# THE DEVELOPMENT OF CHILDREN'S INTERPRETATION OF EMOTION IN MUSIC

LAUREN GAIL WILD



A thesis submitted to the Department of Psychology, University of Cape Town, in fulfilment of the requirements for the degree of Master of Arts

1996

# ACKNOWLEDGEMENTS

I would like to thank my supervisor, Richard Oxtoby, for his encouragement and his insightful but non-intrusive guidance.

My gratitude is also extended to the following people:

- Gillian Finchilescu, for her helpful comments on the statistical analyses

- the teachers at Westcott Primary School and Monterey Pre-primary school who good-naturedly allowed me to disrupt their classes; their pupils, who participated with such enthusiasm, and all the adult volunteers who gave so generously of their time and efforts.

- my family and friends whose patience and practical and emotional support while I have been writing this thesis have been invaluable.

The financial assistance of the Centre for Science Development (HSRC, South Africa) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at are those of the author and are not necessarily to be attributed to the CSD.

# CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# ABSTRACT

The relationship between music and emotion is complex, and has defied explanation for a great many years. The present study addressed one particular aspect of this issue: to what extent can young children interpret the emotions expressed in short musical excerpts drawn from larger works, and how do their interpretations differ from those of adults and change as they mature?

Following an appraisal of current theoretical approaches and empirical research, a structural model of emotions incorporating the ideas of fuzzy, prototypically organised emotion concepts based on underlying dimensions of degree of pleasure and arousal or activation (Bullock & Russell, 1984, 1986; Russell, 1989) was used to reveal and interpret patterns and developmental trends in children's understanding of emotion in music. 5-year-old, 7-year-old and 9-year-old children and adults (n = 30 in each age group) participated in the main study. They were asked to link 18 musical excerpts to an emotion word/facial expression pair selected from the following alternatives: calm, happy, excited, scared, angry/cross and sad. These were presented to the subjects in a set of three tasks utilising different combinations of musical excerpts and emotion words.

The results revealed three main trends in the development of children's interpretation of emotion. Firstly, even preschool children used emotion terms

quite systematically to refer to musical excerpts. These judgments were linked to formal structural/acoustic elements of the musical excerpts, but did not show any consistent relationship with a variety of personal and experiential characteristics of the listener. Secondly, inter- and intra-individual consistency in the labelling of excerpts increased with age. Thirdly, agreement among individuals was found to vary according to the emotion being expressed, and certain emotion words were systematically 'confused' with each other.

These developmental trends are discussed in terms of narrowing emotion categories and a corresponding stabilising of category boundaries between emotion states similar in degree of activation and, to a slightly lesser extent, pleasure or displeasure. Together with changes in the understanding of one or two particular concepts and in the base rates for using certain emotion labels, this decreasing fuzziness of emotion concepts is proposed to account in large part for the developmental advances that have been observed in children's interpretation of emotion in music.

# INTRODUCTION

> Why was it necessary with the old silent movies to hire a pianist to bang out a crudely appropriate musical background? Why must the movies and TV have musical scores?... Why, when I first saw the Grand Canyon and the Piazza San Marco and the Alps, did I feel that these things had all been more moving in cinerama? Why? Because both God and Man forgot to put in the music.
>
> (Brown, 1981, p.240)

At least since the days of Plato and Aristotle, two intriguing questions have continued to perplex scholars of human thought and behaviour. The first of these concerns the nature and functions of music in human experience. The second concerns the nature and functions of emotion.

According to ethnomusicologists music is, and has been, a part of every known culture. Although musical forms and styles vary from one society to another, making and listening to music appear to be fundamental, universal human activities (Blacking, 1987). Children begin singing as soon as they gain sufficient control over their voice to talk (Dowling & Harwood, 1986; Gardner, 1983). And thanks to the availability of electronically reproduced music, the average twentieth century child typically grows up being exposed to a wide range of musical styles in a great diversity of contexts. Most households possess a radio, television, and/or tape-playing equipment, and music is commonly heard in shops, churches, schools, restaurants and discotheques as well as in more formal settings such as concert halls. As Merriam (1964, in Radocy & Boyle, 1979) points out, there is probably no other human cultural activity which is so ubiquitous and which reaches into so much of human behaviour.

Yet in spite of its pervasiveness, music is something of an enigma. Not only are its origins unknown, but it is unclear why it has remained so important in human experience. Music - a phenomenon defined by Sessions as "controlled movement of sound in time" (quoted in Gardner, 1983, p.105) - has no clear adaptive value and is not essential for biological survival. Why, then, is it so difficult to imagine a world without it? What is music *for*?

This question has puzzled some of history's greatest scholars including Aristotle and Darwin, and is unlikely to have a simple answer. A brief consideration of music's functions, however - in public and religious ceremonies, education and therapy, commercial settings, film and television, consumer marketing and advertising - as well as its many informal uses by individuals in our society - suggests that music's ubiquity may be due at least in part to its capacity to heighten our emotional life. The idea that the arts have a special relation to feeling is apparently pervasive in all cultures (Storr, 1992), and there is little doubt that music can arouse strong emotions in people and influence their mood and behavioural state. Furthermore, music is often argued to be able to express, display or represent emotions in all their complexity, and for some it is the capacity of music to communicate ranges and nuances of affect that cannot be expressed verbally which accounts for its chief value (Gaston, 1968; Langer, 1951; Osborne, 1982). It has been claimed that aesthetic education is the education of feeling in quality, depth and breadth (Reimer, 1989) and that music can be particularly useful in breaking down the psychological isolation of individuals diagnosed as autistic (Rosenfeld, 1985), as well as for putting ordinary people in touch with feelings from which they have become alienated (Sloboda, 1985).

Explaining the links between music and emotion therefore appears to be one of the most important and fascinating tasks facing those interested in the psychology of music. This is particularly the case given the current renaissance of interest in emotion within contemporary psychology. After neglecting the study of emotions for many years,

psychologists have gradually come to accept that they are complex and important phenomena which suffuse everything we are and do: an understanding of emotion is therefore a vital component of our understanding of ourselves and others. In fact, Wellman, Harris, Banerjee and Sinclair (1995) have argued that developing an understanding of emotion is an integral and crucial part of the child's emergence into the social-psychological world - a contention supported by recent studies which demonstrate a relationship between children's understanding of emotion and their positive perception of their early school social experiences, their popularity, and their moral sensibility (Dunn, 1995; Manstead, 1995).

Why, then, has an explanation of music's emotional expressiveness and import been largely lacking in the work of music psychologists to date? One possible reason is that despite the current resurgence of interest in emotions, they continue to defy explanation: there is not even general agreement about what an emotion *is*. Another possible reason for the paucity of scientific work on music and emotion is that musical meaning is grounded at least in part in long-standing cultural conventions and deeply rooted beliefs about the nature and functions of art. Because our responses to music partly reflect deeply embedded and unconscious assumptions about the way the world is, questions about it are difficult to pose and to resolve (Francés, 1958/1988). Yet as Berlyne (1972) argues, the same nervous system is used for both aesthetic and other activities, so that if there seems to be an inexplicable gap between aesthetic behaviour and other behaviour, this points to limitations in our understanding of how the nervous system works in non-aesthetic contexts. Deepening our understanding of the meaning, effects and appeal of music must therefore have much to tell us about human nature in general. And other areas of psychology must have much to tell us about the nature of music.

What, then, can people's interpretation of emotion in music tell us about their understanding of emotion in other contexts? And what can our knowledge of their understanding of emotion contribute to our understanding of music? Until now, research on emotions and studies of musical expression have tended to progress independently.

The aim of this dissertation is to draw these two separate threads together in attempting to address one particular issue: how children's understanding of emotion in music develops as they mature. Can young children reliably identify happy music, sad music, calm music, and so on? How does their knowledge compare with that of adults? And to what extent is their recognition of emotion in music dependent on, or influenced by, their experience with different musical styles and genres, their ability to perceive various structural features of music, and/or their understanding of emotions and emotional life more generally? These are just some of the questions that will be considered in the following chapters.

The first four chapters of this thesis provide a general review of the theoretical and research literature concerning people's understanding of emotional expression in music, and examine the current state of knowledge regarding how this understanding develops.

Chapter 1 begins by introducing the main philosophical approaches to the question of emotional 'meaning' in music, and proceeds to examine the contribution of empirical research to answering the question of whether music conveys emotions in ways that can be reliably recognised by listeners.

Chapter 2 takes a step back, as it were, to consider the question of what we actually mean by the term 'emotion': what is the nature of the phenomenon that music is supposedly expressing? A brief discussion of the opacity of the term 'emotion' as it is typically used in psychological theory and research is followed by an outline of the primary theoretical approaches that underlie much of the work in this field. Two contrasting interpretations of emotion, the biosocial or psychoevolutionary perspective and the social constructionist approach are summarised, and their limitations for informing our understanding of emotional expression in music are discussed. This is followed by an argument in favour of shifting the focus of research away from studying emotion states to focusing on emotion *concepts*, and an alternative theoretical approach describing emotion concepts as fuzzy categories organised around prototypical scripts and based on underlying bipolar

dimensions of degree of pleasure and degree of activation or arousal (Russell, 1980, 1991b) is presented. The chapter ends with an exploration of the potential of such an approach to clarify our understanding of emotional expression in music.

Chapter 3 focuses on the role that musical and listener characteristics play in contributing towards music's emotional meaning. The chapter begins with a summary of the empirical literature concerning the relationship between the formal, structural elements of music and the emotions that are perceived as being expressed, and goes on to consider the strengths and weaknesses of a variety of theories of musical expressiveness in accounting for these findings. The question of the effects of personal and experiential characteristics of the listener on the perception of these elements is then introduced.

The suggestion that certain skills or knowledge might be required in order to interpret the emotions expressed in music is developed further in chapter 4, which explores the following question: when do children first become able to interpret emotions in music, and do their interpretations become more reliable with age? Evidence concerning developmental trends in the perception of emotion in music are considered within the perspective of current knowledge about children's understanding of musical elements as well as their knowledge of emotions. The limited utility of traditional theories for explaining these findings is discussed, and an alternative model of emotional development incorporating the idea of decreasing breadth and fuzziness of emotion categories (Bullock & Russell, 1984, 1986; Russell & Bullock 1985, 1986) is explored. Despite having some limitations, this model has illuminated many of psychologists' observations about children's interpretation of emotion words and facial expressions, and it will be used as a framework for the empirical part of the dissertation.

Chapter 5 describes the rationale and method of a research project based on this model, which attempts to fill in some of the gaps in our understanding of the development of children's interpretation of emotion in music. The results of this investigation are presented in chapter 6, and discussed in chapter 7. The final chapter summarises the

main conclusions drawn from the theoretical review and empirical work, and considers their potential implications for the use and importance of music in contemporary society.

There can be few more puzzling questions facing psychologists than that of the relationship between music and emotion, and it is easy to dismiss it as inexplicable by rational means and an unsuitable topic for scientific study. If we do so, however, we are inevitably limiting our understanding of human psychology. The anthropologist Levi-Strauss (in Gardner, 1983) is by no means alone in claiming that a failure to take music seriously weakens any account of the human condition, and the same can be said of emotion. This study represents an exploratory attempt to link music psychology with a broader theory of emotional development, with the aim of contributing to a greater understanding of human cognition and affect. It is believed that such an approach is necessary in order to begin to synthesise many unrelated bits and pieces of data into a coherent framework - a framework which will benefit the research community, and from which it will be possible to draw in order for music psychology to make a more positive contribution towards applied areas as well.

# CHAPTER 1:
# DOES MUSIC EXPRESS EMOTION?

The relationship between music and emotion is complex. Music has long been recognised as a powerful generator and moderator of emotions, although the feelings that it evokes may be difficult to describe, let alone to account for. But music is also inextricably linked with emotion in another way. Not only do listeners often have emotional reactions to music, but music can also *express* emotions in ways that can be recognised by listeners:

> It seems both an unaccountable, and yet, to date, irrefutable fact that one of the correct and most pervasive ways to describe pure music - that is, music without text, title, or program - is in terms of the 'garden-variety' emotions: anger, hope, sadness, happiness, and others of that general kind. (Kivy, 1989, p.153)

Here we are faced with what may be the most perplexing characteristic of music from a psychological point of view. In physical terms, music is merely a combination of sounds of different pitches, timbres, durations, and so on. Why do we attribute emotional significance to these sounds? And when we interpret music in emotional terms, are we hearing what the composer intended to express, or are we interpreting the meaning of the notes in the light of our own background and experiences? Do others hear those notes as we do, or as the composer did? Does the music mean the same thing to everybody?

There are, in fact, those who deny that music can be correctly described in emotional terms. Professional composers themselves have debated whether or not music can be expressive, contradicting each other and occasionally themselves in their attempts to deal with the issue of musical meaning. Whether music does express emotion is clearly a huge and complicated question; nevertheless, it is a question which has captured the interest of aestheticians and philosophers as well as a few psychologists,

and some of the most influential views and findings that they have produced will be outlined below.

## 1.1. Philosophical Theories

The traditional approach to examining the question of expression and meaning in music is speculative or philosophical. While many divergent views have been offered, aesthetic theories concerning the association of meaning with music have traditionally been grouped into two main theoretical positions. The essential debate between these two positions centres around whether the meaning of music is primarily intellectual or primarily emotional.

The *expressionists*, on the one hand, argue that meaning in music is chiefly emotional. Emotion is in some way embedded in the music, and in perception of the music we perceive the embodied feeling. In its simplest form, expressionism holds that the composer is expressing his or her own personal, specific (and perhaps especially deep and insightful) emotions in his or her music; these are transmitted to the audience by a performer who acts as the composer's mouthpiece (Budd, 1989). This view has a long and involved history, and is implicit in attempts to 'explain' a piece of music's mood in terms of the contemporaneous tragic or joyful events in the life of the composer. It has often been criticised, however, as it implies that music is merely a vehicle for self-expression. In addition, it suggests that a composition cannot be sad, for instance, unless its composer was sad at the time and intended to express that sadness (Barwell, 1986). And as Langer (1942) argues, how could a composer possibly feel all those emotions a musical work may seem to express?

There is a more sophisticated version of the expressionist thesis, however, which argues not that music expresses the emotions of the composer or performer, but rather that music in some way represents, displays, expresses or symbolises 'universal' human emotions as understood by the individual composer (Callen, 1982). In other words, the music expresses the composer's knowledge of human emotions, not how or

when that knowledge was acquired. According to this view, music's most valuable and significant feature is its ability to express ranges and nuances of feeling that could be expressed only inadequately, if at all, through verbal means, and which would otherwise remain outside the listener's experience (see, for example, Gaston, 1968 and Merriam, 1964 in Radocy & Boyle, 1979).

In opposition to the expressionists, the *formalists* deny that emotions are the subject matter that music is intended to represent. They argue that meaning in music is primarily intellectual, based on the perception, understanding and appreciation of the formal structural relationships contained within a musical composition (Budd, 1985; Hargreaves, 1986). Thus, musical works have no meaning outside themselves, and an understanding and appreciation of music requires no familiarity with human emotions, or anything from life outside of music (Bell, 1914 in Reimer, 1989).

One of the most forceful proponents of this view is Eduard Hanslick (1891/1986), a Viennese music critic of the late nineteenth century. Perhaps as a reaction against the romanticised, emotionalised interpretations of artworks which abounded at that time, as well as the corresponding belief that 'good' music is music that expresses socially desirable emotions, Hanslick denied that music is capable of representing a specific feeling or emotional state. He argued that definite emotions such as hope, sadness and love involve, and are distinguished from one another by, concepts, judgments, thoughts, and a historical content which music cannot represent. Music may, he argues, embody the *motion* or *dynamics* of feeling, but it cannot depict the feeling itself: "Music can, in fact, whisper, rage, and rustle. But love and anger occur only within our hearts" (Hanslick, 1891/1986, p.9). The formalists do not claim that people are incapable of hearing emotion in music, and at times even describe music in emotional terms themselves. However, they argue that this is not the way music *ought* to be appreciated: emotion terms cannot be used to describe music in a way that is purely musical, and hence are irrelevant to the composition's meaning and significance.

In between these two extremes lies the position taken by Susanne Langer (1942). Like the expressionists, Langer argues that the function of musical works *is* to symbolise feelings or emotions, but she agrees with Hanslick that it is the dynamic form, rather than the specific and complete content, of emotion that music represents. Music symbolises the morphology of feelings, the 'shells' of emotive life, things like tension and release, acceleration and retardation, continuity and sudden change, rather than specific, verbalizable emotions such as yearning or gaiety. Because different emotions sometimes share common dynamic forms, this explains why musical expression is ambiguous, as well as why music can express the nature of feelings in a way that language cannot: "Because the forms of feeling are much more congruent with musical forms than with the forms of language, music can *reveal* the nature of feelings with a detail and truth that language cannot approach." (Langer, 1942, p. 191)

There are many contemporary variations on and developments of the three classical viewpoints outlined above. To some extent, these arguments are based on a particular understanding and use of certain terms such as 'represent', 'symbolise' and 'express', the meanings of which are debatable. When this is taken into account, it appears that the various viewpoints are not in fact as different as they might at first seem. Hanslick (1891/1986), for example, claims that to *represent* "is to produce a clear and distinct content, to put it before our very eyes" (p.14). Scruton (1993) argues that if music is to represent emotions, it must tell a story about them which is capable of being translated into another medium, while Kivy (1989) adds that if music is a representation of certain emotions, the composer must consciously have intended his or her music to portray these emotions, which is often not the case. This conception of representation is central to the formalist argument: Hanslick does not deny that music can *symbolise* or *resemble* emotions, in the sense that a relationship between the two is established in our imaginative interpretations of music; his main point, however, is that emotion is not a characteristic of the music itself. Nevertheless, others have argued that conventional denotation is not the only possible kind of meaning, and that simply because music does not articulate emotions in the way that literature does - because it is difficult to say what a piece of music means - one does not need to go to the extreme of denying that music can be expressive (see, for

example, Langer, 1942). We do, however, need to explain what we mean when we say that music expresses emotion. Kivy (1989), for example, argues that music is not an *expression* of emotion, for the word implies that an inner state is being manifested externally. However, music *expresses* emotion in the sense that it resembles expressive behaviour: music is 'sad', for example, in the same way that a St Bernard's face is 'sad'; it possesses expressive *properties*, but not expressive *content*.

Of course, the differences between conflicting viewpoints are more than just terminological. However, some writers have argued that the dispute between formalists and expressionists is in fact nonsensical: all music offers both formal structure and the possibility of semantic and expressive interpretations, and the affective and intellectual responses to music cannot really be separated. According to this view, confusion and controversy arise only because different scholars - due to their disposition or training - are focusing on different aspects of the same experience (Francés, 1958/1988; Storr, 1992). The musicologist L.B. Meyer, for instance, has argued that there are two types of meaning in music: *embodied* meaning and *designative* meaning. Designative meaning is that which refers outside the music to non-musical objects and events. Embodied meaning, in contrast, is the meaning that a musical passage has for a listener in terms of its own structure and the interaction of that structure with the listener's musical knowledge and expectations, based on past experiences with music of the given style (1956, 1973 in Sloboda, 1985). Thus although some music has quite explicit designative (extra-musical) meaning, this designative meaning does not exhaust its significance - the music will still 'make sense' even if this reference is not appreciated by a listener. On the other hand, no music contains only embodied meaning, for its elements will always possess connotations derived from the musical culture (Cazden, 1951 in Farnsworth, 1969).

Looked at from this perspective, the formalists' and expressionists' views are complementary rather than contradictory. Why, then, the continuing - and often heated - debate? Perhaps part of the answer lies in the fact that these philosophical positions are concerned not only with the meaning of music as such, but also with its value, beauty and importance. Questions of musical meaning inevitably overlap with

some of the fundamental and perplexing issues touched on in the introduction to this dissertation: What is the essence of music? Why does it exist? What is the nature and value of the experience of perceiving and understanding a musical work?

Although the philosophical approach has not given us any clear answers to these questions, several contemporary theorists agree that it no longer seems plausible to define music as essentially an art of emotional expression. Budd (1989) points out that a great deal of fine music does not appear to express emotion and resists any attempt to impose an emotional meaning upon it[1]. There is little support for claims that the highest function of music as an art is the expression of emotion, or that music has some kind of special mission with regard to emotional expression over and above the other arts (Kivy, 1989). Nor is the emotive description of music the only correct and illuminating way of describing it: it is possible to understand a musical composition without reference to any emotional terms - without reference, in fact, to any verbal terms at all - and any explanation of music's value and meaning should not separate it from the medium of music itself (Brown, 1981; Budd, 1989).

But while music should not be seen merely as a vehicle for the expression and communication of psychological states, that it is capable of expressing emotion is still thought to be an important aspect of music. People who care about music want to communicate their experiences to others, and while emotive descriptions of music are not the only way of talking about it, they are without a doubt one of the most common. And although centuries of philosophical debate have not resolved the question of whether it is correct to describe music in emotional terms, lay people and musicians alike apparently find it quite natural to talk about music using words such as 'happy', 'nostalgic' or 'sad', and continue to do so. The next section , therefore, will bypass these great abstract issues and go instead right to the level of the everyday language that people use to communicate about music.

[1] An extreme example is serial music, where the music's 'grammar' - it's embodied meaning - becomes virtually all-important (Budd, 1989).

## 1.2. Empirical Findings

While the traditional philosophical approach has made a valuable contribution to our understanding of musical meaning, it is limited by a lack of objectivity. Approximately fifty years ago, however, psychologists began to take a more scientific approach to the question of whether music can really communicate emotions. Do people agree on the emotions expressed by different kinds of music? And are such emotions pronounced to the extent that it is possible to translate them into such discrete terms as hope, joy and anger?

Empirical work in music psychology is notoriously difficult due to the problems involved in controlling conditions. Studies of musical expression are often characterised by a lack of consistency and clarity in the selection and use of musical stimuli, particularly when it comes to determining what emotions a musical work 'really' expresses. Some researchers have selected existing compositions and used program notes, autobiographies, personal documents, letters and so on to determine the intentions of composers to convey certain emotions. Others have used the judgments of musical experts, of the experimenters themselves, or of a panel of judges; still others have selected instrumental versions of vocal music based on the emotional meaning of the words. Alternative approaches have been to use original works incorporating specific musical elements that experimental research has suggested convey particular emotions, or to make use of improvisations. In addition, the techniques used to measure listeners' understanding of the music vary and provide different degrees of guidance: while the primary approach used by researchers requires listeners to associate verbal responses to music, methods range from the use of adjective checklists through semantic differentials to non-verbal judgments of synonymy.

In spite of these methodological inconsistencies, however, there is considerable agreement among studies that listeners drawn from roughly the same subculture show widespread consistency in their characterisation of the emotional content of Western tonal music, even if they are not highly knowledgeable about the specific kinds of

music in question, and even if the excerpts are taken out of the context of a larger composition and therefore presumably express incomplete musical ideas.

In a pioneering study, Hevner (1936) presented 52 college students with five pieces of music of widely differing character, although all drawn from classical music in the Western tonal tradition and played either by a solo piano or orchestra. She asked them to respond by checking all the adjectives from a list of 66 which they found appropriate to describe each piece. Hevner then developed an 'adjective circle' grouping these adjectives into eight clusters, each cluster containing adjectives with similar meanings and emotional connotations. For example, merry, joyous and gay fell into one cluster; dreamy, yielding and tender into another. The groups were arranged in such a way that as one progressed around the circle the mood similarity would decrease steadily until the opposite cluster was reached; from there back to the starting cluster the resemblance would increase again. Hevner's results revealed a general consistency among subjects in the adjectives checked for all but one of the musical selections. Farnsworth (1969) subsequently updated Hevner's procedures using factor analysis. He revised Hevner's list, including only fifty of the original adjectives and rearranging them into ten adjective clusters reported to possess more mood consistency.

Although these techniques seem fairly valid and reliable, they have not given rise to a great deal of subsequent research. Nevertheless, several researchers have found that subjects tend to agree with each other, with the intentions of the composer or performer, or with the opinions of experts, at above-chance levels when they are asked to link excerpts of classical music with an emotion chosen from a list of between three and thirty adjectives (Campbell, 1942; Hampton, 1945; Meerum Terwogt & van Grinsven, 1991; Thompson & Robitaille, 1992), to choose between a pair of alternative labels for different excerpts (Scherer, 1979), to describe musical excerpts on semantic differential scales (Edmonston, 1966), or to judge the similarity of emotions expressed by different musical excerpts (Brown, 1981)[2]. Agreement

[2] Only studies in which subjects were specifically asked to identify the emotions expressed in or conveyed by the music were included in this review. Although studies in which subjects are asked

between listeners does vary depending on the selection played, and is typically highest for selections which can be characterised as program music. However, the ability to recognise emotions in music does not depend on the ability to identify a particular composition or place it in a context (Hampton, 1945). Music may not, however, be able to express all emotions, or to express every emotion equally well. Mull (1949) and Lowry (1974, in Haack, 1980) focused on humour in music, and found that music can express humour in ways that can be reliably recognised by listeners. Hampton (1945) found that emotional expressions characterised by an unpleasant feeling tone were better identified than emotional expressions characterised by a pleasant feeling tone, and that emotional expressions characterised by what he referred to as a "mental attitude" (such as determination, triumph or defiance) were easier to identify in musical compositions than were the "purely emotional" expressions. Campbell (1942), on the other hand, found that while college students agreed with each other consistently in identifying the emotions of joy, gaiety and assertion in music, there was less agreement for the emotions of yearning, calmness, sorrow and tenderness. More recently, Meerum Terwogt and van Grinsven (1991) and Scherer and Oshinsky (1977) found that happiness and sadness were easier to recognise in music than fear and anger.

These findings are not necessarily universal. Most of the above studies have been conducted with college students, and the extent to which they can be generalised will only be established by further studies investigating other populations and using a methodological rigour sometimes lacking in the early work. The generalizability of the above findings may also be limited to certain musical genres. Not only are the vast majority of studies in this field rooted in the tradition of Western (mainly European) tonal music, but they have tended to deal almost exclusively with serious 'art' music, which is of interest to only a minority of people. Folk and popular music and music from other traditions has been largely neglected, as have musical sound

---

"How does the music make you feel?" have often been included in reviews of the literature on music's emotional expressiveness (see, for example, Radocy & Boyle, 1979), this approach was felt to be problematic: while the emotions expressed by music *may* be the same as those invoked in the listener, this is by no means always the case.

sources, styles, and forms (e.g. electronic music) which have developed since most of the basic work was done in this area.

Another qualification of these findings is that these studies are answering the question of whether music *can* communicate emotions, not whether it typically *does* communicate emotional meaning or information. We do not know whether the listeners in the studies mentioned above, responding spontaneously, would have classified the music in terms of emotional categories. There is, in fact, some doubt as to the efficacy of words in gauging such a complex emotional experience as music presents. Francés (1958/1988) points out that the significations of music, as perceived by individual listeners, do not necessarily have verbal equivalents. Thus, there is not always a possibility of translation. Verbal labels can describe the shifting emotional nuances expressed by music only in a crude and approximate fashion, and it may in fact be impossible to express in words much of what we experience and recognise in music. To quote Imberty (1975, cited in Dowling & Harwood, 1986, p.207): "It is exactly the *meaning* that when explicated in words gets lost among the verbal significations - too precise and too literal - and gets betrayed." As Tighe and Dowling (1993) point out, our knowledge of musical meaning is procedural, rather than declarative. Just as we perceive an angry face directly and do not have to think about or verbalise what emotion the face is expressing, so we perceive meaningful, interpreted musical events, and normally do not have to cogitate about or express in words the emotional meaning we perceive in the music.

Nevertheless, although words could not describe music to someone who has never heard it, they are necessary if we are to be able to discuss and converse about it. And the empirical evidence does suggest that people find it quite easy and natural to attach general emotional labels to pieces of music when they are asked to do so, and that they are quite consistent in their descriptions of music's emotional character. Of course, results indicating that at least some music does possess emotional expressiveness for at least some listeners tell us little more than is already provided by common sense. What these findings do *not* tell us is what it actually means to say that we hear a certain emotion in a particular musical work. What is the nature of the

experience of hearing a sequence of inanimate sounds as being expressive of an emotion?

Any attempt to answer this question has, in fact, to consider several related issues. The first of these centres around the meaning of the term 'emotion'. When we say that music expresses emotion, what is the nature of the phenomenon that is supposedly being expressed? Secondly, if music does express emotion, *how* does it do so? And thirdly, what characteristics does a listener need to possess in order to be able to perceive or understand this expression? These questions will be the focus of the following chapters.

# CHAPTER 2:
# WHAT ARE EMOTIONS? TRADITIONAL THEORIES AND A NEW APPROACH

Underlying the arguments of formalists such as Hanslick is the implicit belief that linking music to emotion somehow demeans music. This belief can be explained in part by a long-standing notion in Western society that emotions are simple, involuntary, disorganised and non-cognitive phenomena which we are seized by, suffer or undergo. 'Emotion' is a fairly new word which first appeared in English and French around the end of the sixteenth century. It replaced the 'passions', those wild, potentially dangerous aspects of human psychology that threaten to corrupt human reason - the faculty that differentiates us from the animals - and to disrupt behaviour, and which continually need to be controlled and suppressed (Crespo, 1986; Lyons, 1992). According to this traditional view, emotion is a lower function than thought, and it is considered with distrust and contempt.

Contemporary psychology has, however, brought us to the realisation that emotion and thought are not absolutely distinct domains, but are inextricably linked. Furthermore, it is now recognised that emotions are complex and important phenomena which play a central role in individual experience and interpersonal relations; in fact, it is difficult to conceive of human life without emotion. As Bullock and Russell (1986) argue, the ability to function effectively in the social world, to interpret each other's actions and transient affective states and guide our behaviour accordingly, is dependent on knowing how to detect and distinguish between different emotions, what the causes and consequences of particular emotions typically are, what the temporal sequence of emotional reactions can be, how to control the experience and expression of emotions, and so on. As Izard (1991) points out, emotions are so important to human existence that it is hard to believe that their scientific study remained outside the mainstream of psychology for more than a century. Indeed, it is only in the last decade or two that we have seen a resurgence of interest in emotions.

In this time, however, the study of emotions has become one of psychology's most rapidly growing areas, attracting a great deal of attention and interest from investigators in clinical, developmental, personality, physiological and social psychology.

Perhaps partly as a consequence of the earlier devaluation and disparagement of emotions, however, the meaning of the term 'emotion' is opaque and ambiguous, and some of the confusion and controversy in discussions of music's expressiveness seem to stem from its vague and inconsistent use. While everyone believes that they know what such words as anger, fear, and happiness mean, no one has yet found a commonly acceptable definition for emotion. Some writers define it as a feeling, others as a cognitive state or type of judgment, others as a physiological response, a brain state, or a set of behaviours, and still others as a social process (Beckwith, 1991; Izard, 1991). While illustrations can be found to support each definition, each can also be met with criticisms and counterexamples.

Consequently, many psychologists believe that attempts to define emotions in strictly cognitive, strictly behavioural, strictly physiological or strictly social terms are automatically inadequate because emotions are multidimensional constructs that include all these and other components (Brody, 1985; Fehr & Russell, 1984). Thus for Izard (1991), "an emotion is a complex phenomenon having neural, motor-expressive, and experiential components" (p.54). As Fehr and Russell (1984) argue, however, such a definition is so broad that almost any psychological process would qualify as an emotion, and Izard himself has admitted its inadequacy. Fehr and Russell observe that many writers less inclined towards philosophical argument have therefore simply ignored the problem, presumably considering emotion to be a self-evident concept which needs no definition. But this leaves unanswered important questions such as which words in the English language actually refer to emotions. Bullock and Russell (1986) point out that most people will probably agree that anger, happiness and fear are emotions, but what of admiration, courage and loneliness? How many emotions are there? What is an adequate operational definition of emotion? Are facial, vocal and gestural behaviours all alternative measures of the

same construct? If an infant smiles or frowns, can we conclude that he or she is happy or angry? And can we legitimately talk about emotions in our pet dogs, in fish, in insects or - to return to the particular focus of this thesis - in abstract stimuli like music and the visual arts?

Underlying many of these questions is the issue of whether emotions are primarily biological or cultural in nature. Or, to put it slightly differently (cf. Wierzbicka, 1992a), do emotion terms provide labels for emotions that already have an independent existence, or do they contribute to the construction of emotions themselves by creating certain assumptions, norms and expectations which may influence the interpretation of emotional experience, guide behaviour, and shape interpersonal relationships?

## 2.1. Theories of emotion: Psychoevolution versus social constructionism

The biosocial or psychoevolutionary interpretation of emotions has a long history, and is clearly expressed in Darwin's (1904) theory of emotions as common to animals and humans, based on primitive states of physiological arousal, automatically triggered, and manifested in innate and universal behavioural responses and expressions. According to this perspective, emotions have organising and motivational properties that serve adaptive functions, and evolved in order to promote the chances for survival of the individual and the species.

Theorists who support the biosocial position typically take a theoretical perspective which argues that there is a set of discrete, fundamental and universal human emotions. These fundamental or basic emotions are thought to comprise specific neural substrates, distinct, automatic patterns of physiological expression and particular phenomenological experiences which are innately linked. They are usually assumed to correspond to basic emotion terms of the English language such as joy or jealousy (e.g. Ekman, 1993; Izard, 1991; Plutchik, 1962). Today, $7 \pm 2$ basic emotions are most often listed: anger, fear, happiness/joy/enjoyment, sadness, and disgust, with surprise, contempt, shame or shyness and interest sometimes added.

According to this *discrete-emotions* model, all emotion words other than the basic 7 ± 2 can be defined in terms of the fundamental emotions. Either they are seen as 'blends' of these emotions (e.g. Plutchik, 1962), or they are seen as subdivisions of or variations on these basic emotions which are grouped together in 'emotion families' of related states (Ekman, 1993). Cognitive and sociocultural processes may play a role in determining the intensity of emotions, in distinguishing the different members of emotion families, or in determining what will trigger an emotion, as well as in providing 'display rules' which determine how and whether an emotion will be expressed. However, these processes are seen as peripheral to the essence of emotion.

In contrast to the psychoevolutionary emphasis on a universal set of biologically-based emotions, the social constructionists argue that emotions are "the internal representation of social norms or rules" (Averill, 1986, p.100). In other words, emotions are learnt, culturally variable patterns of behaviour which express a particular culture's conception of what is appropriate behaviour in a specific, socially defined situation (Griffiths, 1989). According to this view, there are no natural, fixed and non-arbitrary units of emotion called things like 'happiness' which objectively exist as separate categories of experience. Rather, 'emotion' itself is a socially constructed concept, and instead of forming coherent bundles, any pattern of physiological changes, subjective experience, expressive reactions and behaviours is as likely to occur as any other (Bullock & Russell, 1986): no single component or subset of components represents a necessary or sufficient condition for any emotion. While some of these component responses, for example certain expressive reactions, may have a biological basis, the way in which these components are organised into coherent sets of responses is culture-bound and determined primarily by social interaction rather than biological evolution (Averill, 1986; Gordon, 1989).

Evidence can be found for both of the positions discussed above. However, differences in theoretical position on this issue are reflected in clear differences in methodological strategy, and the data presented by each does tend to focus on different topics and only rarely to overlap. The biological theorists, for example, cite evidence that certain facial (and to some extent vocal and gestural) expressions of

emotions are cross-culturally recognised: in other words, the percentage of judges selecting the intended emotion term from a list of alternatives in labelling a particular emotional expression is significantly greater than chance, and greater than some higher arbitrary level (usually 70%), across all cultures. They have also presented evidence of cross-cultural agreement about which facial expressions fit different social situations such as the death of a child, a fight, or seeing friends (Campos & Barrett, 1984; Ekman, 1993; Ekman & Friesen, 1971; Ekman et al., 1987; Matsumoto, 1992; Mesquita & Frijda, 1992; Van Bezooijen, Otto & Heenan, 1983). The social constructionists, on the other hand, argue that agreement about the meaning of facial expressions using a forced choice procedure implies cross-cultural *similarity*, but not necessarily cross-cultural *equivalence,* in the interpretations given to these expressions, and point out the less than perfect agreement among people in the labels given to faces, particularly in cultures speaking languages that are not of Indo-European origin. As evidence for their alternative view, they refer to studies showing that even some of the most important (and supposedly fundamental) English emotion words have no equivalent in some languages. Polish, for instance, does not have a word corresponding exactly to the English word 'disgust' (Hiatt, 1978, reported in Russell, 1991a). And many emotions that some cultures would regard as equally fundamental are either omitted from lists of basic emotions in English, or have no equivalent in that language (Kagan, 1984; Mesquita & Frijda, 1992; Russell, 1989, 1991a; Wierzbicka, 1986, 1992b).

If we look at the evidence more objectively, however, it becomes clear that it does not support either approach to the exclusion of the other. Reviews by Mesquita and Frijda (1992) and Russell (1991a) suggest that there are both cross-cultural differences and cross-cultural similarities in the various subcomponents of emotion, and that similarities in one aspect (for instance, cognitive appraisal) do not necessarily imply similarities in other aspects (such as behavioural expressions). Therefore while the biological and constructionist views appear to be diametrically opposed and are usually considered to be rivals, some theorists have argued that viewing emotion from only one side of the debate narrows our understanding of the concept, and that these two approaches need to be integrated rather than set in opposition to one another (e.g.

Harris, 1989; Lutz & White, 1986). Epstein (1984) and Griffiths (1989) have both attempted to do this by arguing that different emotions may in fact be different types of psychological phenomena. Epstein, for instance, poses a distinction between primary and secondary emotions, where primary emotions such as sadness or fear are biologically determined and adaptive patterns of physiological arousal, feeling states and expressive reactions, and secondary emotions such as jealousy are culturally determined syndromes that have no characteristic elements of physiological arousal, feeling states or expressive reactions. He points out, however, that this distinction is theoretically useful rather than practically useful: in everyday life the distinction between primary and secondary emotions becomes blurred because the primary emotions have often been shaped by learning to the extent that they are functionally indistinguishable from secondary emotions. Such theories have the potential to prove enlightening, but they need to be further developed before they can offer a thorough and succinct explanation of emotions.

The next question that needs to be asked is what these debates contribute to our understanding of emotions as expressed in music. The answer seems to be very little, apart from raising questions about whether what is communicated in music are phenomena characterised by a distinctive universal signal and for which there exists a correct label such as 'happy' or 'sad', or whether music expresses an undifferentiated emotional domain which we divide up and interpret according to certain assumptions, norms, expectations and verbal labels provided by our culture. Although cross-cultural research on the emotional meaning of music is virtually non-existent, an early study by Gundlach (1932) does suggest that although there are similarities in the way in which different cultures express and interpret emotions in music there are, not surprisingly, also marked differences.

## 2.2. A new approach: Studying emotion concepts

If questions about what exactly emotions are had to be clearly answered before we could investigate emotions in music, research in this field would probably never progress very far. A more useful approach, however, has been provided by those

psychologists who have recently begun to shift their attention away from emotion *states* to focus on emotion *concepts.* Rather than asking questions such as what emotions exist or what the nature of the phenomenon that is labelled by terms such as 'anger' *is*, they are concerned with investigating how people use emotion terms within the perspective of ordinary language to refer to people, situations, and so on (see, for example, Bullock & Russell, 1986; Fehr & Russell, 1984; Harré, 1986; Lutz & White, 1986; Russell, 1989, 1991a; Storm & Storm, 1987; Wierzbicka, 1992a, 1992b).

Gowin (1970, in Reimer, 1987, p.80) describes a concept as

> a sign which points to a commonality in events and which permits the concept user to make relatively stable responses to those varied events. The signs which are vehicles for the concept are largely linguistic and conventional. The commonality in events may range from simple similarities to regularities to law-like invariance.

A concept therefore has three essential features. Firstly, it requires a phenomenon that is manifested more than once: a common feature, or set of features, of a range of events. Secondly, it needs a linguistic or conventional sign, symbol, name or marker of some sort which indicates the common feature(s) being noticed. The third thing needed for a concept to exist is that this symbol should be regularly and stably associated with the phenomenon to which it refers, even though the phenomenon exists in various forms incorporating a range of common qualities. Thus according to this definition, emotion concepts such as love or anger comprise words which categorise a number of experiences or phenomena which are somewhat related to each other.

People's emotion concepts undoubtedly reflect the underlying emotion domain - their emotion experiences - to some extent at least. However, they also reflect the implicit, unexamined common-sense assumptions of a particular culture (including its scientists and philosophers), linguistic factors such as the language's history, and practical issues involved in the need to communicate about emotions (Wierzbicka, 1986). Leaving aside, for now, the question of how, or to what extent, emotion

concepts contribute to the construction of our experiences and social realities, it can be argued that human emotion (as a basic psychological reality or experience) and human emotion concepts do not necessarily have a one-to-one correspondence. The relationships between categories and phenomena are unclear and multiple, and the fact that a particular language does not have a word for a certain emotion does not tell us that the speakers of this language cannot perceive that emotion as a distinct, recognisable feeling or experience, or that they cannot talk about it. What it does suggest, however, is that this culture does not recognise the emotion in question as particularly salient (Mesquita & Frijda, 1992; Wierzbicka, 1986, 1992a).

According to this view, everyday emotion concepts are worthy of study in their own right. Thus the ambiguity of the term 'emotion' is not a problem, but rather a data base on which to found an investigation of the different and changing ways in which emotion terms are used in everyday life. No formal definition of emotion is required because, in a sense, the definition of emotion, and of particular emotion states by people in a particular society, *is* the variable under study (Crespo, 1986; Masters & Carlson, 1984). So from this perspective, rather than trying to find out what 'it' is that is communicated by happy music, we should be concerned with the way in which people use emotion terms such as happiness, yearning, and so on to categorise and label music. We can probably never determine whether a piece of music is *really* sad, but we can ask on what basis the word 'sad' is used by people in contemporary society to refer to music and, in turn, what this use tells us about their conceptions of and assumptions about sadness (as well as about music).

### 2.2.1. Prototypes and fuzzy categories

One consequence of the shift of attention from emotion *events* to emotion *concepts* is that some researchers have begun to question our ideas about how emotion concepts can be defined. According to Bullock and Russell (1984), it has traditionally been assumed that natural language concepts about everyday events can be defined in the

same way that logically defined concepts or 'proper sets' are: as clear, nonoverlapping entities whose membership is defined in terms of a set of necessary and sufficient features. Every item possessing these criterial attributes therefore has a full and equal degree of membership in the category (Rosch, 1975). Gradually, however, this classical or Aristotelian view of concepts has begun to be challenged as evidence has been accumulated to show that many natural language concepts or categories such as vegetable, vehicle or furniture lack clearly defining or criterial features. The same problem apparently applies to the concept of emotion: nobody has yet been able to identify the necessary and sufficient conditions for applying the term.

Consequently, Fehr and Russell (1984) argue that we need to look for an alternative way of defining emotion concepts, and suggest Rosch's (1975, 1977) prototype theory as the most likely candidate. Rosch's categorisation of natural language categories, which develops an idea initially suggested by Wittgenstein (1953, in McCloskey & Glucksberg, 1978) has already been very useful for reconceptualising psychological concepts like intelligence, mental illness and specific diagnostic categories (Fehr & Russell, 1984). Although it is not the only alternative to classical definitions of emotion concepts, Rosch's theory has helped to explain the difficulties psychologists have experienced in defining emotion, suggested new directions for research, and revealed interesting properties of emotion concepts previously overlooked. In particular, it has provided a framework for understanding how each emotion word categorises a huge realm of possibilities of subjective experience which nevertheless have something in common.

According to Rosch's theory, systems or taxonomies can be viewed as both vertically and horizontally organised. Firstly, category systems are vertically organised within an abstract-to-concrete hierarchy, typically involving three main levels: the *superordinate* (furniture, or emotion); the *basic* (chair, or anger), and the *subordinate* (kitchen chair, or annoyance). The basic or middle-level categories (usually represented by the shortest words) are those that people use for much of their everyday conversation and thought, because they offer the best compromise between

informativeness and cognitive economy or simplicity[1] (Shaver, Schwartz, Kirson & O'Connor, 1987).

More importantly for our purposes, however, Rosch's (1975) prototype theory proposes that each of these levels is also horizontally organised. This horizontal organisation has two main (related) features: firstly, natural language categories have an *internal structure*, or hierarchy of inclusion, and secondly, natural language categories have *fuzzy* boundaries. Russell (1989) explains the meaning of these features as follows: (1) Membership in natural language categories is organised around typical or characteristic members known as *prototypes* (the clearest cases or best examples). (2) Membership within a category is a matter of degree rather than all or none, based on family resemblance to the prototype. Good members have many attributes in common with the prototype, while poor members have less psychological similarity to the prototype. (3) Borders between categories are vague, rather than clear-cut, and categories tend to overlap each other (in ways that vary in kind and number) rather than to be mutually exclusive. Some categories overlap each other completely, others to a high degree, others to a minimal degree, and some not at all. (4) Although some items clearly belong in one category, others are difficult to categorise and are inconsistently said to be or not be exemplars of overlapping concepts.

Evidence that the superordinate (English) concept of emotion has an internal structure (i.e. that membership in the concept is a matter of degree rather than all or none) and fuzzy boundaries (i.e. that no sharp boundary separates members from nonmembers such as states of mind, moods, affective disorders and bodily sensations) has come from several converging sources in a set of studies by Bullock and Russell (1986), Fehr and Russell (1984), Fehr, Russell and Ward (1982), Russell (1980, 1983), and Shaver et al. (1987). For instance, these studies have shown that although subjects have difficulty giving an explicit rule-like definition for emotion, they are able to

[1] The term 'basic' as it is used here has a slightly different meaning to that given it by the biosocial theorists, who might argue that some basic (i.e. middle-level) emotion concepts are more basic (i.e. fundamental) than others.

make reliable judgments about whether particular words represent good or poor members of the category. Limited cross-cultural support for a prototype analysis of the concepts most closely resembling emotion in other languages has also been obtained by Smith and Tkel-Sbal (1995) and Smith and Smith (1995).

In addition, evidence has been provided to suggest that basic or middle-level emotion words themselves, such as anger and happiness, can be thought of as labels for fuzzy sets with an internal structure. In other words, it is argued that a term such as 'anger' can be defined as a category lacking sharp boundaries into which actual occurrences of various emotional states fit more or less well, and which overlaps with other categories such as disgust and sadness. For example, Russell and Bullock (1986) examined how adults categorise the message conveyed by emotional facial expressions, and found that some expressions were prototypical examples while others were intermediate or poor examples. There were also borderline cases in which subjects could not decide whether particular facial expressions were or were not members of a particular category.

As Fehr and Russell (1984) point out, however, this evidence should not be interpreted to imply that prototype analyses of everyday concepts are *better* than classical definitions, nor that classical definitions have no role to play in the study of emotions. It may be that for each emotion there is a correct word to describe it, but that people simply have less than perfect knowledge about the criteria for perfect membership (see, for example, Wierzbicka, 1992a). McCloskey and Glucksberg (1978) have made the point, however, that if we are concerned not with developing a technical taxonomy for emotion terms but rather with determining people's *representations* of emotion categories as currently used in everyday life, our focus needs to be not whether classical definitions of emotion are possible, but rather whether people act as if they are clearly defined. And the existing data seem to indicate that they do not.

### 2.2.2. Emotion scripts

If emotion categories are organised around prototypes, we need to ask what the nature of those prototypes is. One view which is gaining increasing acceptance is that the concept expressed by an emotion word is a *script* or *paradigm scenario* (De Sousa in Beckwith, 1991; Fehr & Russell, 1984; Lewis, 1989; Russell, 1989, 1991a; Shaver et al., 1989; Wierzbicka, 1992a). According to this position, an emotion is not a 'thing', but rather a sequence of subevents, and to know the meaning of a word such as 'fear' is to know such a sequence. An emotion script is therefore a knowledge structure which contains prototypical causes, desires, beliefs, physiological reactions, feelings, vocal and facial expressions, self-control procedures, actions, consequences, and so on, which are ordered in a causal sequence. Some of these features are likely to be universal, others specific to a particular culture. We interpret observed or experienced emotions by implicitly comparing them with the prototypical emotion script: the more features an emotion event has in common with a script, the clearer a case it is of that particular emotion. For example, consider the prototypical case of fear as described by Fehr and Russell (1984, p. 482):

> A dangerous situation occurs suddenly. You are startled, and you scream. You try to focus all your attention on the danger, try to figure a way out, but you feel your heart pounding and your limbs trembling. Thoughts race through your mind. Your palms feel cold and wet. There are butterflies in your stomach. You turn and flee....

Actual instances of fear vary in how well they match this prototypical case of the concept fear.

Unfortunately, the word 'script' is somewhat vague and is used differently by different writers. Wierzbicka (1992a), for example, argues that emotion concepts can be defined in terms of prototypical scripts formulated in terms of thoughts, wants and feelings which specify necessary and sufficient conditions for particular emotion concepts, and which draw clear boundaries between apparently synonymous concepts. Her notion of emotion scripts has, however, been criticised for being too narrow and

not conveying anything about what it *feels* like to experience a particular emotion such as sadness.

For these and other reasons, Russell and his colleagues, as well as other theorists such as Averill (1986), use the word 'script' in a broader sense, arguing that the features that make up an emotion script are neither necessary nor sufficient for applying a particular emotion term. An actual emotion event may resemble the script to varying degrees and in different ways. On the one hand, the more features of the prototypical script present, the closer the resemblance and the more appropriate the script label. On the other hand, each feature of the script has a prototypical value, and subevents of an emotion (for example, beliefs or actions) can resemble the prototypical subevent of the script again to varying degrees and in different ways. For example, while researchers have identified a prototypical facial expression of fear, a range of different facial expressions count, to varying degrees, as fear, and the border between this and expressions of other closely related emotions are fuzzy. What remains unclear, however, is just how abstract or concrete are the features that constitute a script, or whether different people - even within the same culture - might possess slightly different scripts for the same emotion (Russell, 1989, 1991a).

An important aspect of the prototype theory of emotion concepts as outlined above is that terms such as fear, anger and sadness have fuzzy boundaries and overlap with one another to a greater or lesser degree. This leads us to the question of how emotion concepts are related: how do people perceive the similarities and differences among emotions? One possibility is that the similarity between different emotion concepts is determined by the *number* of features of a script (cognitive appraisals, feeling states, overt actions, and so on) that they have in common. An alternative possibility is that specific features of a script have different 'weightings' and that some features are more important or fundamental than others in determining the similarities and differences between emotions. The latter view is implied by Russell (1989), who argues that the degree of overlap between different emotion states depends in large part on their position on two underlying bipolar dimensions of degree of pleasure and degree of activation.

### 2.2.3. Dimensions of emotion

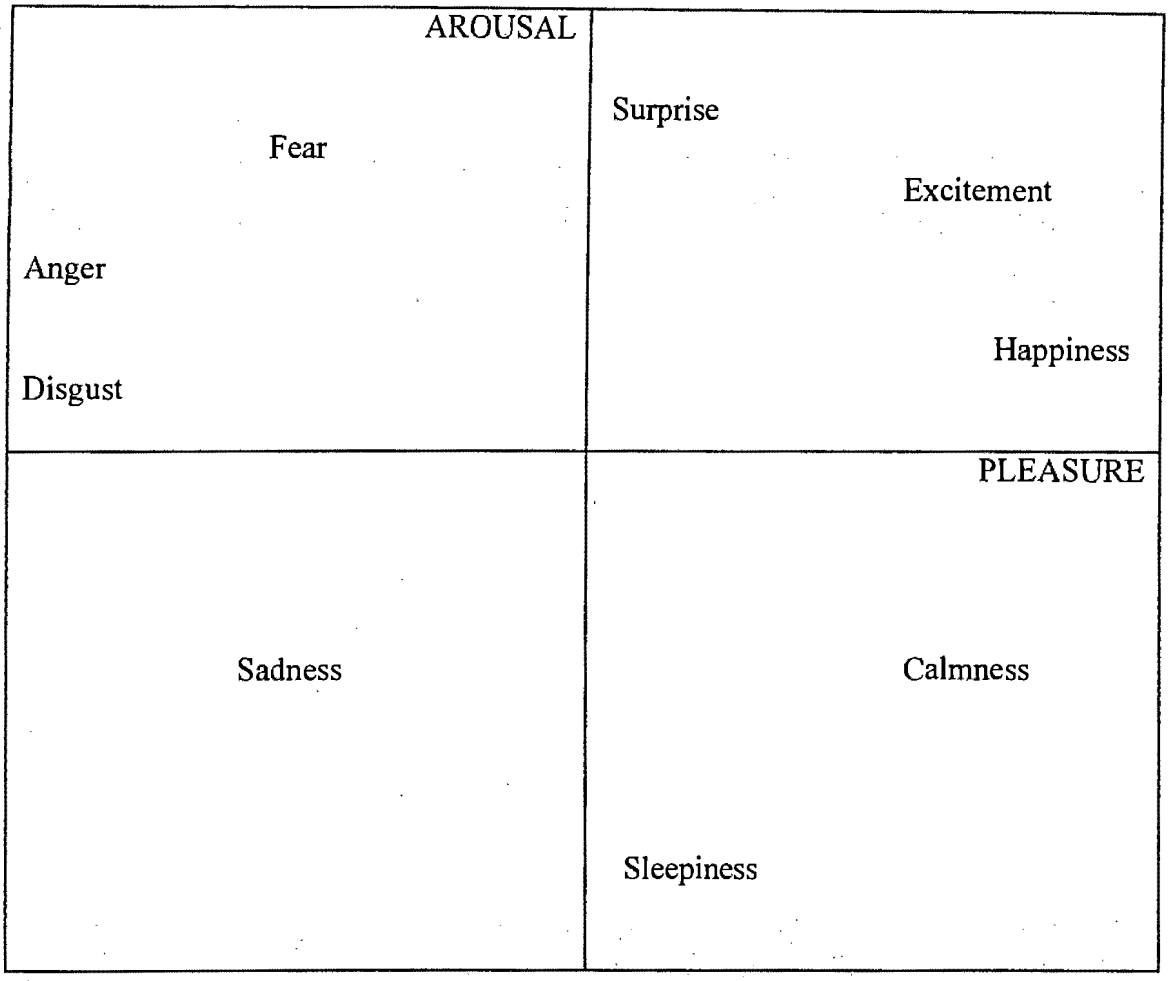
The idea that the domain of emotion is structured - that is, that the categories labelled by words such as 'happy' and 'afraid' may be interrelated in a systematic fashion - was first introduced by Woodworth (1938, in Bullock & Russell, 1984). Woodworth examined the 'errors' people made when categorising facial expressions of emotion, and found that these errors were systematic: subjects who did not choose the correct emotional term for a particular facial expression nonetheless chose from a limited set of categories. Woodworth consequently proposed a structural model of emotions in which emotions are ordered along a continuum in such a way that adjacent emotions are the ones most likely to be confused with each other. In other words, the most common label for one expression is likely to be somewhat applicable to adjacent expressions. Woodworth's student Schlosberg (1952) later modified this model. Noticing that the two ends of Woodworth's continuum were also occasionally confused for one another, Schlosberg placed these two ends together. The result was a model which organised emotions in a circular structure. Schlosberg interpreted this structure as being based on two underlying dimensions which he termed pleasantness - unpleasantness and attention - rejection. Later, he suggested a third dimension labelled sleep - tension (Schlosberg, 1954).

Since Schlosberg's ideas were published, psychologists trying to clarify, understand and organise emotion concepts have attempted to identify the primary dimensions underlying these concepts. Although there is a certain amount of diversity in the terms used by different writers, two main orthogonal dimensions have been frequently noted and are well established. The first of these is typically labelled as pleasure - displeasure, valence, comfort - discomfort or hedonic tone, and the second as degree of activation or arousal - sleep. These dimensions have been identified in adults' and children's categorisation of emotion words and facial expressions as well as in studies of self-reported mood and of the emotional impact of physical environments, using statistical techniques such as multidimensional scaling and factor analysis (Bush, 1973; Davitz, 1969; Moland & Whissell, 1993; Russell, 1978, 1980, 1983, 1989, 1991a; Russell & Bullock, 1986; Shaver et al., 1987; Storm & Storm, 1987; Whissell

& D'Elia, 1993). Furthermore, both of these dimensions have been found in studies using languages other than English, although only the pleasure - displeasure dimension has been consistently supported cross-culturally (Lutz, 1986; Russell, 1991; Russell, Lewicka & Niit, 1989). Russell (1989) has suggested that these dimensions might represent pancultural aspects of emotion scripts, whereas other features are culture-specific. However, it is possible that these dimensions might take on slightly different forms in different cultures. Storm and Storm (1987) have suggested that these two dimensions might be manifested slightly differently at different points in the emotion domain, for example 'pleasure' might take on the form of moral goodness or social desirability, and activation might also be interpreted as intensity or degree of arousal. Studies by Lutz (1986) and Smith and Tkel-Sbal (1995) suggest that the same argument might also hold cross-culturally.

The most clearly developed and widely tested model utilising these dimensions in explaining the interrelationships among emotion concepts is the circumplex model of affect developed by Russell and his colleagues (Bullock & Russell, 1984, 1986; Fehr & Russell, 1984; Russell, 1978, 1980, 1983, 1990, 1991a, 1991b; Russell & Bullock, 1986; Russell & Ridgeway, 1983). This model, which is depicted in Figure 1, places emotion words in a roughly circular order around the perimeter of a two-dimensional space, so that more similar emotions fall closer together, while emotions labelled as antonyms fall opposite one another on the circle. Underlying this circular ordering are the two dimensions described above: pleasure versus displeasure on the horizontal axis, and degree of arousal on the vertical axis.

**Figure 1:** Circular structure of emotions in two-dimensional space (after Bullock & Russell, 1984, 1986)




Evidence for other dimensions has also been found, although these typically have accounted for small amounts of variance in the perceived relations, and their identification has been inconsistent. The most commonly found third dimension has been interpreted variously as potency, power, control or dominance, and may refer to the individual's position in relation to other persons in the emotion-eliciting situation (Davitz, 1969; Lutz, 1986; Shaver et al., 1987), and/or to the depth or profundity versus superficiality or frivolity of the emotions (Averill, 1975 in Russell, 1978). While researchers are just beginning to go beyond this third dimension, there is some suggestion that dimensions beyond pleasure-displeasure and arousal may refer to cognitive appraisals or beliefs about the antecedents or consequences of the emotion, such as locus of causation (Russell, 1978).

An advantage of the dimensional approach to emotion concepts is that it allows for the possibility that the complete set of emotion words in any language underdifferentiates the actual emotion domain, and goes at least some way toward accounting for how different emotions such as anger, fear and hate overlap and fuse with each other. Critics of this approach have, however, argued that the emergence of particular dimensions depends on the particular sample of words used in a study, and that at face value some of the terms used in dimensional studies (for example, 'sleepy') probably do not denote emotions at all (Scherer, 1984). Although there is some evidence that almost perfect-looking circumplex analyses such as Russell's require the inclusion of words that most people would not call emotions, a study by Shaver et al. (1987) has shown that a similar structure does emerge even for emotional terms which have not been preselected to fit a circumplex - or any other theoretical - model, although parts of the circle are not occupied. Nevertheless, there are still those who argue quite persuasively that two or three dimensions are unlikely to be able to explain the variation in meaning conveyed by the large set of emotion-related terms in a language such as English.

Few, if any, dimensional theorists would in fact deny this limitation (see, for example, Russell, 1978, 1991b; Schlosberg, 1952). Rather, both dimensional theorists and those who argue against this approach have suggested that dimensions such as pleasure - displeasure and degree of activation are most suited to describe the internal feeling component of emotions and its expression in observable reactions, but that they do not allow for finer discriminations among emotional states which rely on other components of emotion such as antecedents and consequences which are more culture-based (Davitz, 1969; Russell, 1978; Scherer, 1984; Schlosberg, 1952). In addition, it needs to be remembered that we are dealing with emotion *concepts*, not emotion *events*. The categorisation of emotions may reflect emotion 'reality', but it may not, and talking about emotions in terms of dimensions or structures does not necessarily - as has often been assumed - compete or contrast with talking about emotions in terms of categories. Rather, the two approaches are complementary. Bullock and Russell (1986) point out that if we consider emotion categories to have fuzzy boundaries and to overlap one another, asking in what way categories overlap

leads directly to the idea of a structure according to which emotion concepts are interrelated. And the idea of intercategory structure in turn leads directly to the idea of underlying dimensions. From this perspective, categories (based around prototypical emotion scripts), structures and dimensions are all interrelated aspects of the network of conceptions used by humans to understand emotion.

### 2.2.4. Music and the prototype analysis of emotions

How can we apply this model to our understanding of the emotional meaning of music? Firstly, a prototype analysis of emotion categories would suggest that a piece of music is said to be sad, happy, or whatever to the extent that it resembles certain ideal or prototypical cases, with certain musical works being good examples of a particular emotion category, and others being poorer examples of that category. Furthermore, if the categories overlap one another and are not clearly bounded, then a particular piece of music could be a peripheral member of one emotion category and a prototypical member of another. A single piece of music may share certain features in common with several emotion categories, and therefore be categorisable as expressing more than one emotion.

In addition, the applicability of a certain emotion concept or concepts to a particular piece of music should depend on those concepts' positions on one or more underlying dimensions. Since it is widely agreed that musical compositions (or, for that matter, decontextualised facial, vocal, or bodily expressions of emotion) are incapable of expressing the cognitive or interpersonal aspects of emotion such as antecedents, appraisals, causes, or behavioural consequences, the most important dimensions are likely to be those of pleasure-displeasure and degree of activation, with some contribution possibly also made by a potency/depth dimension. Frijda (1958, in Davitz, 1964) has suggested that perhaps facial cues express the more general characteristics of emotional expression (for instance, positive versus negative, active versus passive), while situational cues determine the more specific and subtle discriminations within a general category (such as anger versus impatience, satisfaction versus contentment), and it seems reasonable to assume that the same

might apply to music. Thus according to Russell's circumplex model, the emotion terms most applicable to describing a particular piece of music could be predicted by their co-ordinates on the dimensions of pleasure-displeasure and degree of arousal, with the applicability of emotion terms for describing that piece of music decreasing as one moves around the circumference of the circumplex model. In addition, we would expect that if people's categorisation of music is primarily determined by these underlying dimensions, they will not be able to make as fine discriminations between the emotions expressed in music as they would in other contexts where additional information about causes, consequences and concomitants specified by the emotion script occur. As a result, all emotion terms similar in pleasantness and arousal might be equally applicable to describe a particular piece of music.

Although Russell's circumplex model has not been examined directly in relation to musical expression, there is some empirical evidence to suggest that it might be useful for exploring emotional meaning in music. Meerum Terwogt and van Grinsven (1991) asked thirty adults to rate how well music is capable of expressing each of 100 emotions. They found that happiness, anger and sadness were rated as best expressed in music, followed by calm, restful, and restless. These results were interpreted to suggest that music expresses both a positive-negative and an activity dimension. The authors subsequently found that when subjects were asked to match musical excerpts with emotion terms, they made a fairly systematic distinction between music which was positive or negative in tone, and that within the negatively valenced pieces there was clear consensus about the identity of 'sad' music (which is low in activation) as opposed to music expressing anger or fear (both high activation states), which tended to be confused with each other. This supports the findings of McMullen (1974, in Abeles, 1980) who examined similarities between the Hevner checklist approach and the dimensions of the semantic differential (SD). In his study, college students were asked to describe eight classical and pop music excerpts on 7-point scales using adjectives that had either previously been shown to be important in SD research or that were representative of each of Hevner's eight adjective clusters. The results suggested that the Hevner adjectives are describing two sub-factors of the evaluative dimension produced by most SD research, and also hinted at the presence of an

activity dimension. Another relevant finding is that of Hampton (1945), who found that emotional expressions characterised by either a clearly unpleasant or a clearly pleasant feeling tone (for example, grief or joy) were easier to identify than emotional expressions where the feeling tone was subdued or neutral (such as abandon, praise, or resignation).

Wedin (1972) took a different approach. She applied factor analysis and multidimensional scaling to listeners' ratings of musical excerpts ranging from Bach to the Beatles on 125 adjectives, and found that emotions expressed by music fall on three dimensions which she termed *intensity - softness* (contrasting energy, intensity, power and activity with softness, relaxation, tenderness, intimacy, and sometimes melancholy); *pleasantness - unpleasantness* (contrasting joy, mirth, happiness, elation, gaiety and pleasure with agony, fear, depression, threat, uneasiness, tension, gloom and sorrow); and *solemnity - triviality* (contrasting solemnity, dignity and grandiosity with exhilaration, popularity and triviality). In a smaller-scale study, Asada and Ohgushi (1991) asked thirteen music students to rate the 18 pieces of Ravel's *Bolero* on semantic differential scales and found that two (unnamed) dimensions emerged. The first contrasted words such as energetic, powerful, open and gorgeous with calm, quiet, suppressed and sober, resembling Wedin's 'intensity' and Meerum Terwogt and van Grinsven's 'activity' dimensions. The second dimension contrasted terms such as comic, sticky and coquettish with words such as serious, flowing, and reserved, showing some similarity to Wedin's 'solemnity-triviality' dimension.

There is also some indication that certain dimensions may be easier to use in interpreting music than others. Meerum Terwogt and van Grinsven (1991) found that disagreements were more likely to occur on the pleasure dimension than on the activity dimension: for example, all subjects identified a Tchaikovsky extract as passive, but there was some disagreement about the happy-sad distinction. Similarly, Campbell (1942) found that her subjects confused the 'active' emotions of gaiety, joy and assertion with each other, and also tended to confuse the 'passive' emotions of yearning, tenderness, calm and sorrow; in contrast, they had no difficulty in

differentiating between two pleasant, but differentially active emotions like gaiety and calmness. Likewise, Scherer and Oshinsky (1977), in a study utilising synthesised tone sequences, found that subjects tended to agree more readily on the activity level than on the dimensions of pleasure and potency.

Further indirect support for the utility of a prototype analysis of emotion in music comes from an early study by Campbell (1942). She found that whereas subjects tended to agree among themselves when it came to dividing musical pieces into seven main (middle- or basic-level) emotion categories, there was little agreement among them for subdivisions within each of these main categories. For example, while over 90% of her sample agreed that certain selections expressed gaiety, there was no agreement as to whether this gaiety was carefree, lively, or merry. Campbell comments, however, that it is unclear from her results whether such distinctions are not expressible in music, or whether her subjects lacked the musical sensitivity to distinguish them.

What these studies do not tell us, however, is what the features are that categorise prototypical musical examples of particular emotion categories. It is all very well to say that a piece of music will tend to be described as sad if it conveys displeasure and low arousal, but this does not in fact tell us anything more about *how* it is that music expresses these emotional qualities. According to Radocy and Boyle (1979), variables contributing to musical meaning can be divided into two main categories: (1) those related to the structural features of the music itself, and (2) those related to the listener. The following chapter will examine each of these in turn.

# CHAPTER 3:
# HOW DOES MUSIC EXPRESS EMOTION?

## 3.1. What musical elements express emotion?

A number of researchers over the years have attempted to link the emotional properties of music to specific acoustic elements and their combinations. Scherer and Oshinsky (1977) found that two-thirds to three-quarters of the variance in listeners' judgments of the emotions expressed in synthesised tone sequences could be explained by the manipulation of musical elements. However, no definitive taxonomy of musical elements used to express emotion has yet been developed. Definitions vary, and what is a chief component in one taxonomy is often only a subcomponent in another (Bruner, 1990). Part of the reason for this seems to be that different studies have focused on different musical elements, and not controlled the elements that others have used as independent variables. Edmonston (1966) and Eagle (1971, in Radocy & Boyle, 1979), for example, found that rhythm seems to be primary in determining the emotional character of music. Henkin (1955, 1957) found that rhythmic and melodic factors were equally important in determining the emotional meanings of music, while Hevner (1937), Scherer (1979) and Scherer and Oshinsky (1977) found that tempo had the most influence on judges' ratings. Hevner (1936, 1937) also found, however, that the relative importance of musical variables depended on the emotion that was being expressed.

Some of these differences between studies may be due to methodological features. While psychological studies investigating the *manner* in which music expresses recognisable emotions often show greater methodological sophistication than those which merely investigate the extent of agreement among listeners about music's emotional character, they do differ in their choice of music, the techniques they use to measure listeners' responses, and their methods of statistical analysis. Researchers also differ in the extent to which they take a *synthetic* (experimental) versus an

*analytic* (naturalistic) approach. In the former, particular musical variables are isolated and manipulated and their effects on the listener's responses measured. In the latter approach, researchers measure people's interpretations of music taken from 'real life' (Hargreaves, 1986; Radocy & Boyle, 1979), which has an advantage over the former approach in ecological validity, but involves a corresponding loss of precision. In addition, while the pioneering researchers in this area had to rely on a skilled musician varying one formal property of a piece while keeping the others roughly constant, the advent of computers has led to the use of electronically synthesised tone sequences in which musical elements can be systematically and independently manipulated. In spite of these methodological variations, however, there is some agreement about the affective meaning conveyed by certain musical characteristics, which will be summarised under the headings of time-, melody-, harmony-, modality- and texture-related findings. A glossary of the musical terms used is provided in Appendix 7.

Time-related findings

The two main time-related aspects of music are tempo and rhythm. As far as the former is concerned, researchers have arrived at the same general conclusion: all other things being equal, faster tempos tend to be associated with greater happiness, excitement and activity than slow tempos (Eagle, 1971 in Radocy & Boyle, 1979; Gundlach, 1935; Hevner, 1937; Nielzén & Cesarec, 1982; Rigg, 1940; Scherer, 1979; Scherer & Oshinsky, 1977; Wedin, 1972). Two main aspects of rhythm have been studied: (a) strength and smoothness, and (b) phrasing and articulation (the actual length of notes). Firm rhythms were associated by Hevner's (1936) subjects with words such as dignified and vigorous and by Wedin's (1972) subjects with words such as dramatic and ominous. In both studies flowing rhythms were associated with descriptions such as happy and playful. Some cross-cultural support for these findings is provided by Gundlach's (1932) analysis of Indian music in which war songs were found to have the least even rhythms, and healing songs to have the most even rhythms. Gundlach (1935) also found that European musical phrases with uneven rhythms were classified as delicate, dignified, sombre, and so on, while

phrases with many smooth rhythms were described as brilliant, animated, flippant and glad. In contrast, however, Nielzén and Cesarec (1982) found judgments of gaiety as opposed to gloom to be associated with a marked rhythm.

A relationship has also been found between staccato-legato articulation and the activity dimension of music. Gundlach (1932) and Wedin (1972) found that staccato rhythmic articulation was associated with greater energy and excitement, while Nielzén and Cesarec (1982) found that staccato music expressed more happiness and activity than legato music.

Melody-related findings

Three main aspects of melody have been investigated: mean pitch level, pitch range, and the direction of the melodic line. Scherer (1979) found a high pitch level related to judgments of happiness and surprise, and low pitch levels to disgust and boredom. A high pitch level was found by Hevner (1936) to reflect humour and sprightliness and by Wedin (1972) and Heinlein (1928) to reflect emotions such as happiness (gladness) and playfulness. Similarly, Swinchoski (1947, in Haack, 1980) found that higher pitches were associated with humorous, graceful, happy moods. In all four studies a low pitch level was found to be associated with negative and unpleasant or gloomy emotions as well as with qualities of dignity and solemnity.

As far as pitch range is concerned, Scherer (1979) and Scherer and Oshinsky (1977) found small to moderate pitch variation to be associated with unpleasant emotions such as disgust, anger, fear and boredom, while extreme pitch variation was associated with pleasant, active and potent emotions such as happiness and surprise. Similarly, Gundlach (1935) found that pieces with a pitch range of more than an octave were perceived as more uneasy, animated, grotesque, brilliant and glad, while those with a narrow range were judged as more tranquil, dignified, delicate, mournful, awkward and sombre.

The findings concerning the direction of the melodic line are less consistent. Gundlach (1935) found no relationship between melodic line and the emotions listeners ascribed to music. Hevner (1936) found some tendency for an ascending melody to express dignity and a descending melody to seem exciting, but these effects were very slight. Scherer and Oshinsky (1977), on the other hand, found that rising contours were associated with potent emotions such as interest, surprise, anger and fear, while down contours were associated with boredom, pleasantness, and sadness. Trehub, Cohen and Guerriero (1987, in Trehub, 1993) found that ascending 9-note sequences typically led to 'happy' judgments, while descending sequences led to 'sad' judgments.

Harmony-related findings

Harmony has received somewhat less attention by researchers than the melodic aspects of music. However, Hevner (1936), Nielzén and Cesarec (1982) and Wedin (1972) all found that consonant harmonies were associated with pleasant emotions such as happiness, playfulness, serenity and peacefulness, while dissonant harmonies were associated with descriptions such as exciting, dramatic, agitating, ominous, and sad. Winold (in Farnsworth, 1969) found music with consonant harmonies to be generally characterised as dignified, spiritual, triumphant or majestic. Slightly dissonant music tended to be described as calm, dreamy, happy, and so on, while extremely dissonant music was characterised as depressing, frustrated, humorous, agitated, exciting, and so on.

Modality-related findings

Several studies have tested conventional thinking in Western culture that the major mode expresses happiness and the minor mode sadness. Trehub, Cohen and Guerriero (1987, in Trehub, 1993) found that only musically trained subjects made the conventional major-happy/minor-sad associations when listening to short melodic sequences. Similar results were obtained in an early study conducted by Heinlein (1928), who found that the ability to associate major chords with happiness and minor

chords with sadness was dependent on musical training. Crowder (1984), however, has reanalysed Heinlein's data and found generous support overall, some exceptions notwithstanding, for the conventional connotations for both musically trained and untrained subjects. The effects of musical training were barely significant. Other researchers (Crowder, 1985; Hevner, 1935; Nielzén & Cesarec, 1982; Scherer & Oshinsky, 1977; Wedin, 1972) have found the expected associations of the major mode with pleasant emotions such as happy, playful and peaceful, and of the minor mode with negative emotions such as sadness, gloominess, tension and anger regardless of musical training, although Hevner (1935) did find that these effects were stronger for musically trained subjects. Hevner also found that modality was not related to excited-calm ratings.

Texture-related findings

Two main aspects of musical texture have been examined: volume or intensity, and instrumentation or timbre (tone quality or 'colour'). In an early study, Gundlach (1935) found that the loudest pieces tended to be described as triumphant or animated, and the softest as delicate or tranquil. More recently, Asada and Ohgushi (1991) found that loud music tended to be characterised as energetic and powerful, while Nielzén and Cesarec (1982) found that louder music was judged as expressing more gaiety and activity than soft music. Wedin (1972), on the other hand, found that volume tended to interact with tempo, so that loud music tended to be described as active and energetic when it was played at a fast tempo, while loud, slow music tended to be described as dignified and solemn. Scherer and Oshinsky (1977) studied the variation in volume rather than volume level per se, and found that a large variation was associated with active and potent emotions such as anger and fear, while a small variation in volume was associated with happiness, pleasantness and activity. While little attention has been paid to instrumentation, Gundlach (1935) found that brass instruments carried the melody in songs described as triumphant or grotesque; woodwinds expressed uneasy, mournful and/or awkward feelings; melodies on a piano were perceived as delicate, sentimental and tranquil as well as brilliant; while string sounds were characterised as glad. Behrens and Green (1993) found that

subjects' accuracy at identifying the emotion that a musician intended to express depended both on the instrument played and on the expressed emotion. Subjects were significantly more accurate when identifying sad improvisations performed on a violin or vocally than on a trumpet or timpani, but were also more accurate at identifying sad trumpet improvisations than scared or angry improvisations played on this instrument. However, for angry improvisations they were correct more often when these were performed on a timpani rather than on other instruments. Scared improvisations were most accurately identified when performed on the violin. Taking a somewhat different approach, Asada and Ohgushi (1991) found that pieces played by a larger number of instruments were rated as more energetic and powerful than those played by a solo instrument or only a few instruments.

Scherer and Oshinsky (1977) investigated another aspect of timbre. By manipulating the filtration cut-off level of their synthesised tone sequences, they were able to study the effects of the number of harmonics (overtones) on listeners' emotional judgments. Sounds with few overtones (between 1 and 8) were rated as significantly more pleasant, happy and bored, whereas sequences for which more harmonics (16) were audible were judged as significantly more active, potent, angry, disgusted and fearful.

There are, of course, other musical elements which might contribute to a piece's emotional meaning which have not yet been investigated, and there are other aspects of those features which have been investigated which may well be of great importance in the expressiveness of music (Hevner, 1936). Furthermore, while several studies have investigated the *main* effects of musical elements such as tempo, the *interaction* effects have received much less attention. As Francés (1958/1988), Hevner (1937) and Nielzén and Cesarec (1982) observe, a musical work is a complex of many interacting elements that are not usually isolated, and the typical emotional effect of any one element when it occurs by itself may be neutralised, obliterated, emphasised or otherwise changed by deliberate changes in other musical elements. Thus while these studies identify general trends, many exceptions to these trends can be found in the musical literature. Scherer and Oshinsky's (1977) study has suggested that while most of the variance in music's emotional expressiveness can be explained by the

additive combination of main effects, in some cases specific combinations of musical elements serve to communicate different types or subdivisions of a certain emotion such as happiness (for example, quiet, subdued bliss versus buoyant gaiety). Although examination of these interaction effects will result in a huge increase in complexity and require more detailed specifications of emotional states than have been used in the studies to date, exploration of these effects does seem to be an important task for future research.

Keeping these cautions and limitations in mind, it does nevertheless appear that listeners may well use an informal but fairly standard 'syntax' for interpreting emotion in music. Table I presents, for a selection of emotions, the musical elements which have been found to contribute to the variations in listeners' attributions of emotional states.

**Table I.** Musical elements related to the expression of specific emotion states

| | HAPPINESS | SADNESS | EXCITEMENT | CALMNESS/ SERENITY | ANGER | FEAR |
|---|---|---|---|---|---|---|
| TEMPO | fast | slow | fast | slow | fast | fast |
| RHYTHM | flowing, fast | firm slow | — | flowing | firm slow | — |
| ARTICU-LATION | staccato | legato | staccato | legato | — | (stac-cato?) |
| PITCH LEVEL | high | low | — | — | — | high |
| PITCH RANGE | large | small - moderate | large | small - moderate | small - mod. | small - mod. |
| MELODIC CONTOUR | (up?) | (down?) | (down?) | — | (up?) | (up?) |
| MODE | major | minor | — | — | minor | (min.?) |
| HARMONY | simple, consonant | complex, dissonant | complex, dissonant | simple, consonant | dissonant | disso-nant |
| VOLUME | medium - loud | soft | loud | soft | — | — |
| OVERTONES | few | — | many | few? | many | many |

If, as suggested, listeners infer emotions in music based on specific musical elements, the next question we need to ask is, of course, how it is that these elements achieve their effects.

## 3.2. How do musical elements express emotion?

There are two main positions on the question of how emotional meaning is conveyed by music. The first alternative, usually known as *emotivism*, holds that music's expressive qualities are dispositions of the music to arouse corresponding emotions in listeners. The other view, sometimes referred to as *cognitivism*, argues that music possesses emotional qualities as perceived phenomenal properties of itself (Kivy, 1989): we hear the emotion *in* the music, rather than feeling it in ourselves. Each of these positions, in turn, encompasses various subtheses as to how these effects are achieved.

### 3.2.1. Emotivism

According to the emotivist theory of musical meaning, to say that music expresses an emotion (for instance, sadness) is to say that the music typically makes an attentive listener feel that emotion (it is saddening) (Davies, 1991). Although there are various explanations as to how and why music is capable of evoking emotions in listeners, the best known and most widely accepted is that of Leonard Meyer (1956, 1973 cited in Radocy & Boyle, 1979 and Sloboda, 1985). Meyer's explanation of musical meaning is based on the theory that emotion is aroused when an expectation, or tendency to respond, is inhibited or blocked. Due to a listener's knowledge of and previous experiences with a particular musical style, he or she will have certain expectations about what musical events and patterns will follow each other in a work. Some of these expectations will be thwarted, some will be fulfilled at once, and others will be fulfilled after a delay. This gives rise to a dynamic flux of tensions and resolutions which gives meaning to the music and arouses emotion. Because music in a style with which a listener is totally unfamiliar will not produce expectations in

the listener, he or she will therefore find it totally meaningless and without emotional import.

Not all writers feel that Meyer's explanation is sufficient to account for the meaning and effects of music. However, most would accept that music can arouse or modify our emotions. Certainly at times the emotions expressed in a piece of music and the emotions aroused in a listener are difficult to separate, and probably for this reason the emotivist thesis has widespread support. Emotivism has, however, also been widely criticised as being simplistic, illogical and inaccurate (e.g. Kivy, 1989). This theory implies that when we say that music expresses emotions, we are projecting moods or feelings from ourselves to the music because we have become victims of the pathetic fallacy[1] (Budd, 1985). Critics have argued, however, that we can hear music as sad without feeling sad - if this were not the case, we would expect people to avoid listening to sad music. And it seems doubtful that listening to angry music typically makes people angry. Furthermore, a piece of music might make someone feel sad or angry (for example, when it has a personal association for them), although they would not regard that piece as expressing sadness or anger (Kivy, 1989). Outside of aesthetics, people do not usually say that something that makes them feel sad or angry expresses sadness or is itself angry. As Stecker (1984) points out, the term 'expressing' implies showing something forth, and unless emotion is more closely related to the music itself than emotivism hypothesises, it is difficult to see how we could justify talking of music as expressing emotion at all.

### 3.2.2. Cognitivism

The alternative position to emotivism holds that when we say that music expresses an emotion, that emotion is phenomenologically in the music itself, and nowhere else. Philosophers within this tradition generally draw on the language of semiotics to

[1] Ruskin (1856, in Budd, 1985) believed that our emotions can induce a falsity in our experience of the world, so that - because of their effects on us - nonhuman animals and inanimate objects can seem to possess characteristics that we know they are incapable of possessing. If someone who listens to music is so powerfully moved by it that he or she comes to perceive it as something which is in an emotional state, then this is an instance of what he termed the pathetic fallacy.

explain musical expression. According to this view, music is a *sign* which represents (in the broadest sense of the word) emotional states. However, there is disagreement as to how this representation or signification takes place: it has been suggested that musical works 'symbolise', 'resemble', 'imitate', 'refer to', 'denote' or 'stand for' emotions.

A useful approach to exploring the ways in which music can represent or signify emotion is to use the classification of the ways in which signs can represent other things developed by C.S. Peirce (1955) in a series of essays first published around the turn of the century. Peirce described three types of sign: *index*, *icon*, and *symbol*. An *index* represents its referent by having being associated with it in the past, for example, lightning and thunder with a storm. An *icon* represents because it has a formal similarity to its referent (for example, a wiring diagram representing a circuit), or because it stands in a natural relation which allows the sign to suggest what it represents (for example, a guttering candle used in a film to represent the process of dying). *Symbols* represent by being members of a formal symbolic system, such as a language: their meaning depends on the conventional syntactical rules by which they are combined, rather than on association or resemblance.

Music as index

Indexical representation involves the direct association of a musical form with some extramusical object or event, so that emotions previously associated with the extramusical object come to be associated with the music (Dowling & Harwood, 1986). This phenomenon is, of course, well known to psychologists as classical conditioning. The idea here is that the perceived emotion is not intrinsic to the music itself, but becomes attached to the music through a learning process: people within a given cultural group *learn* that music with certain characteristics reflects certain emotions while music with different characteristics expresses other emotions.

There are two main ways in which this associative learning is thought to occur. Firstly, people learn as a cultural norm that certain types of music are conventionally

used on particular occasions. For example, in our musical culture there are particular musical forms traditionally used in activities like worship, battle, affirmation of patriotism, mourning, burlesque, folk and popular dance, and no doubt many others (Callen, 1982). Consequently, whenever the emotional significance of such occasions is more-or-less the same for everyone, we associate the matching type of music with that emotion (Meerum Terwogt & Van Grinsven, 1991). And so we hear grief in a funeral march, happy exuberance in Dixieland jazz, and so on.

Powerful as these effects may be, however, there is nothing particularly *musical* about them. Indexical representation of this type cannot explain *all* instances of emotionally expressive music, nor how people are able to recognise emotions in unfamiliar kinds of music. Indexical representation or classical conditioning may, however, also be used to explain music's expression of emotions at a more general level. According to this view, it is possible that people attribute emotional meanings to musical motifs or sequences through association with the kinds of emotional words that habitually seem to accompany the motifs over a long period of time in songs, masses, operas, and so on. It is argued that while emotions were initially attached to musical forms according to the whims of the composer, eventually these associations became codified, and people heard what they were supposed to hear in the music (Farnsworth, 1969).

An important feature of this theory of music's emotional expressiveness is that in indexical representation, the relationship between sign and signified is arbitrary. Just as the association between the majority of words in a spoken language such as English and their referents is merely conventional - there is nothing about the word 'cat', for example, which makes it particularly suitable to stand for the small feline domestic mammal as opposed to anything else - so the expressiveness of music is simply a function of customary, culturally variable patterns of association of certain musical features with certain emotional ones, quite apart from any structural analogy between them (Cunningham & Sterling, 1988; Kivy, 1989). In terms of this argument, some different pairing of musical structure and emotion would be just as easy to learn and use as those we actually have.

The indexical theory does appear to be the simplest explanation of how a listener brought up within a particular musical culture comes to understand that culture's music as expressing emotions. There is certainly little doubt that correlations between musical elements and particular emotions vary between cultures and over time within the same culture, and that familiarity with a particular musical culture is often necessary for an adequate appreciation of musical expression. For instance, the eighteenth-century music of Bach and Handel was rooted in part in musical conventions which associated certain musical keys with particular emotions. F major, for instance, was the key of the pastoral idyll, while F-sharp major was a transcendental key. Similarly, Northern Indian music claims to be able to bring out certain emotional nuances using certain corresponding musical intervals (Blacking, 1973). Yet few, if any, composers writing outside of these musical traditions would accept that there is an inherent correspondence between these musical features and their emotional connotations.

The indexical theory of musical expression has not, however, generated much interest or discussion. One possible reason for this is that it gives rise to a 'chicken and egg' situation: it does not explain why composers should have begun to pair emotions with particular musical elements in the first place (Cooke, 1959). Furthermore, it is difficult to know how this theory might be tested empirically. Heinlein (1928) argued, based on the number of 'errors' that his subjects made, that the major and minor modes do not have intrinsic emotional characters but rather have arbitrary conventional emotional associations that listeners learn. Support for this idea comes from the fact that in other musical systems (for example, African, Oriental, Spanish, Slav and Balkan folk-music), pleasure is expressed by music of a decidedly minor character (Cooke, 1959). However, several other studies (Crowder, 1984, 1985; Hevner, 1935; Nielzén & Cesarec, 1982; Scherer & Oshinsky, 1977; Wedin, 1972)) have found that the conventional emotional associations of these modes do not depend on musical training. Yet these studies do not in themselves refute the classical conditioning thesis, since the many everyday experiences that even musically 'untrained' listeners have with music may be sufficient for them to learn the conventional emotional connotations.

Whatever the reason, however, the majority of theorists have rejected the notion of intrinsic neutrality for music. Yet while these theorists agree that certain musical features are more suitable than others for suggesting particular emotions, they are divided as to the reason for this.

Music as icon

Philosophers such as Kivy (1989) and Tormey (in Barwell, 1986) have put forward the theory that the expressive properties of music result from resemblances between characteristics of the music and characteristics of the natural behavioural expressions of emotion. In other words, formal intrinsic properties of music such as tempo, dynamics, harmonic texture, melodic contour, timbre and rhythm represent iconically expressive features and behaviour - the rhythms, vocal intonations, movements, gestures and demeanour - that we associate with specific emotions. Furthermore, the processes involved in understanding these relationships are argued to be common to human beings in all societies (Harwood, 1976).

In some musical traditions, for example the songs of the Venda, variations in melody and rhythm are sometimes deliberate attempts to reflect changes in speech tone (Blacking, 1973). Even within European music, however, many of the emotional cues discussed above do seem to be similar to the non-verbal cues used in making inferences about the emotions expressed in speech. Reviews of the research on emotional vocal cues in speech by Scherer (1986) and Sundberg (1982), and individual studies by Costanzo, Markel and Costanzo (1969), Scherer, Banse, Wallbott and Goldbeck (1991) and Williams and Stevens (1972) suggest the following:

TEMPO: Slow tempos are associated with sadness, sorrow and boredom; fast tempos are described as expressing joy, anger and fear as well as neutrality or indifference.

RHYTHM: 'Legato' articulation is associated with the expression of tenderness; 'staccato' articulation with the expression of anger.

PITCH LEVEL: Lower pitches are associated with sadness and boredom; higher pitches help to express joy, anger, fear, and in one study (Costanzo et al., 1969) also grief.

PITCH VARIATION: A small pitch variation is typically found in sad expressions, whereas a large pitch variation characterises happiness, anger, and fear.

VOLUME/INTENSITY: A soft volume is associated with sadness and boredom; a louder volume with joy and anger. The findings for fear have been inconsistent.

As Scherer (1986) points out, most of the existing studies on the vocal communication of emotion have focused on a few tried-and-tested vocal cues that are relatively easy to obtain, and ignored other potentially important features such as voice quality or timbre. Nevertheless, as with music, it has been found that the aspect of emotion best expressed by these vocal cues is the degree of activation or arousal, and that errors in identifying vocal expressions of emotion are generally a function of similarity in levels of activation of the emotions concerned rather than, for example, similarity in hedonic tone (For example, mistaking anger for happiness and sadness for love) (Albas et al., 1976; Davitz, 1964a, 1964b; Scherer, 1986; Van Bezooijen et al., 1983). Scherer (1986) points out, however, that these findings are hardly surprising considering that most of the vocal cues that have been examined are related to the tenseness or laxness of the voice, and they do not provide a sufficient basis to answer the question of whether only the activity dimension is expressed by verbal cues, or whether there are specific patterns of acoustic elements for each emotion.

On the other hand, the correspondence between emotional expression in music and in the human voice is not exact. Scherer et al. (1991) and Burns and Beier (1973) found that in nonverbal speech auditory cues were particularly helpful when assessing anger and least useful when judging happiness (in Dolgin & Adelson, 1990). Music, in contrast, appears to be able to express happiness more easily than anger (Meerum Terwogt & van Grinsven, 1991; Scherer & Oshinsky, 1977), suggesting that there is something in the music itself, apart from or in addition to its similarity to prosodic features of the voice, that is responsible for its emotional significance. In evaluating the iconic theory of musical expression, however, it is interesting to note that there are

fairly substantial similarities between the cues which listeners use to judge emotion in both speech and music.

Further support for this theory comes from a rather different source: studies investigating the relationship between acoustic cues and body movements. Francés and Bruchon-Schweitzer (1983) conducted an unusual study which revealed a partial equivalence between the verbal categories associated with musical excerpts and those associated with bodily expression sequences recorded either after the actors had listened to the musical excerpts or after they had been given the verbal categories associated with the music. In another investigation, Clynes and Nettheim (1982) asked subjects to press on a button-like transducer in such a way as to express one of seven emotions, and transformed these into auditory signals varying on the following parameters: starting pitch, rising or falling pitch contour, pitch range, and intensity. Subjects were then asked to listen to the sounds and to identify the emotions they expressed using a forced-choice procedure. Recognition of emotions was highly significantly better than chance except for love and reverence which were largely confused with each other. Clynes has used these findings to take the iconic thesis one step further and suggest that the dynamic expressive communication of specific emotions in various sensory modes (including music) are based on common underlying brain programs of spatial-temporal patterns which he calls *essentic forms* (Clynes, 1986; Clynes & Nettheim, 1982). Clynes's data, methods and arguments have been questioned and criticised, and Blacking (1987) observes that he neglects the cognitive aspects of musical ideas as well as of emotions. Nevertheless, his ideas are intriguing and have the potential to stimulate interesting discussion and research.

An important aspect of iconic representation in general, however, is that the sign is vague and ambiguous without contextual support (Brown, 1981). Therefore Tormey argues that having a set of formal musical properties is neither a necessary nor a sufficient condition for a work having expressive properties, and that similar musical properties may have a variety of different expressive meanings in the same work, depending on the musical and/or dramatic context (in Stecker, 1984). By analogy, a particular expressive behaviour is not necessarily the normal effect of only one

emotion. Crying, for instance, might be an effect of joy as well as of sadness and despair. Therefore, while the words of a song or the story of an opera or ballet provide sufficient context so that the emotional meaning of the music is usually clear even on a first hearing, without such a context there might be disagreements or misunderstandings as to the intended emotional meaning of a piece of music.

Although iconic representation is currently one of the most popular explanations of music's emotional expressiveness, a limitation of this thesis has been noted by Scruton (1993), who points out that this theory cannot explain why, if the features we notice in expressive music are precisely those we see and hear in human behaviour, we need the special experience of understanding music in order to appreciate them.

Music as symbol

By symbol, Peirce (1955) meant a sign that depends for its meaning and significance on its relationships in a network with other signs. According to one school of thought, musical motifs are symbols in the sense that they express emotions through their place in a musical pattern which is governed by a conventional, formal system of syntactic rules.

Cooke (1959), in his book The Language of Music, has presented one of the most thorough versions of this thesis. He argues that melody is the most important element of Western tonal music since 1400, and that certain intervals of the diatonic scale are more suitable than others for suggesting certain emotions due to the tonal relationships (tensions) that occur between their members. According to his argument music is a language of the emotions, not simply in a vague general sense, but in the sense that it should be possible to draw up a "dictionary" which would specify the emotional meaning of the basic terms of musical "vocabulary". For example, Cooke argues that the major third expresses joy, the minor third grief, stoic acceptance, or tragedy; the minor sixth falling to fifth expresses anguish, while the sequence 13565 in the major is used to express a simple, innocent, blessed joy.

Cooke argues that the correspondences between melodic motifs and different emotions are inherent, implying that the emotional significance of particular musical intervals is the result of fundamental features of human psychology and physiology and should therefore be universal. Sloboda (1985) points out two corollaries of this argument. Firstly, anyone who is able to comprehend the tonal relationships in music and who has a normal understanding of the emotions should, in principle, be able to discover analogical relationships between music and emotion. Thus an explicit word-music association is not necessary to come to appreciate emotional meanings in music (although it may contribute to this process). Secondly, it follows that pairings of melodic motifs with emotions *other* than the ones we now have would be more difficult to learn, since they would be working against the 'ready-made' analogical relationships inherent in the tonal system.

Cooke collected a large array of examples from Plainsong to Stravinsky to support his analyses, and presents an impressive argument. There are also, however, so many exceptions to Cooke's thesis that it has been sharply criticised, and he has been accused of picking his examples to support his case. As Sloboda (1985) points out, however, counter-examples are not as immediately damaging to Cooke's thesis as they may seem, because Cooke himself admits that melody is only one of the features which composers may use to express emotion. He concedes that volume, rhythm, tempo and harmony can emphasise and modify the emotional significance of the melodic elements in a variety of ways not available to spoken language. So a melody would have different emotional connotations when played fortissimo by the entire brass section of a symphony orchestra than when played liltingly by a solo violin.

Although Cooke's thesis is empirical, it is difficult to know how to test it. As Sloboda (1985) points out, the 'ideal' practical demonstration of this thesis would be given by someone who had never been exposed to any word-music associations (through song, etc.), but who had received otherwise normal musical and linguistic experience, showing culturally 'normal' appreciation of the emotional connotations of music. Of course, it is unlikely that such conditions could ever be set up.

Gabriel (1978) did attempt to study Cooke's thesis empirically. He played each of Cooke's basic melodic motifs or 'terms' to 22 non-music polytechnic students twice. On one occasion the motif was paired with Cooke's emotional characterisation of it, on the other by a randomly chosen description from Cooke's set. Subjects were asked to rate the appropriateness of each motif-emotion combination on a five-point scale. Gabriel argued that if particular melodic motifs convey the meanings ascribed to them by Cooke, then subjects should rate Cooke's designated description of each motif as more appropriate than the randomly chosen one. In fact, the results showed that the designated descriptions and the random descriptions were equally often chosen as more appropriate. Thus Gabriel failed to demonstrate any relationship between melodic motifs and emotional connotations.

Sloboda (1985), however, claims that Gabriel's experimental design may have destroyed the conditions under which it is possible for listeners to link tonal patterns with emotional connotations. Gabriel himself admits that the meaning of the motifs would not be at "maximal power" in his experimental design. Sloboda argues that music's emotional meaning depends on the context in which the crucial melody is embedded: we may need to know where an excerpt comes in the structure of the music before we can interpret it emotionally. Because Gabriel presented the motifs in musically 'disembodied' form, his task was an unreal one and does not in itself disprove Cooke's thesis. As Kivy (1989) argues, musical elements which do not themselves express emotion can take on an emotional connotation by means of their 'syntactical' function in a wider musical context. Thus the emotional meaning of a diminished triad alone, for example, is ambiguous. But in a musical context, during a long period in the history of our musical tradition, it has become a chord that tends to go somewhere, to lead to something, and when it occurs as part of a musical 'sentence' it therefore may take on a restless, 'anguished' character. Gabriel's study does suggest, however, that it is highly unlikely that *every* occurrence of a melodic motif has emotional significance. More likely, it is the portion of a sequence that is highlighted in its position in the musical structure that is responsible for much of the music's emotional meaning (Sloboda, 1985).

Nevertheless, there are many who argue that music is not a language in the strict sense of having a vocabulary of separable elements with fixed connotations, which can be combined by means of syntactical rules in such a way that they take on complex connotations without detracting from their individual meanings (e.g. Langer, 1942). The emotional effects of music, it is argued, are more dependent upon context, less upon purely musical devices, than Cooke allowed. For example, Storr (1992) draws attention to the fact that the French carol *Quelle est cette odeur agreable, Bergers qui ravit tous nos sens* appears as a beautifully tender melody when sung as a carol; but the same tune serves as a rumbustious drinking song, *Fill every glass*, in The Beggars Opera. Cooke's thesis is of course also limited from a psychological point of view in that he explicitly confines his discussion to European art music written since 1400 and in the tonal tradition. As Blacking (1973) points out, because Cooke's theory is not general enough to apply to *any* culture or society, it is automatically inadequate as an explanation of the emotional expressiveness of music - even only of European music. For these and other reasons, the view that is currently most widely accepted is that it is the more primitive, iconic features of music that are the principal carriers of musical meaning.

Integration of different views

It is, however, quite possible that music can be involved in more than one sign function at the same time. Peirce (1955) did not intend his categories of signs to be mutually exclusive, and there is an intuitively appealing argument which holds that the indexical, iconic and symbolic theories of musical expressiveness *together* account for how music expresses emotions: sometimes one, sometimes the other, sometimes - inextricably - all bound up together. The musical expression of emotion probably does not always take place in the same way.

The idea that musical expression may be partly explained by iconic resemblances to emotional behaviour and partly explained by conventional associations was put forward some time ago by Francés (1958/1988). More recently, Kivy (1989) has

articulated this argument in a more sophisticated and systematic fashion, suggesting that emotionally expressive cues in music are of (at least) the following kinds:

i. Those that are heard as expressing emotion because they resemble expressive behaviour of some kind, for example the 'weeping' figure of grief [iconic representation].

ii. Those that do not in themselves resemble expressive behaviour but which contribute, in certain contexts, to expressing emotional properties such as tension and release. An example would be the 'restless' diminished triad [symbolic representation].

iii. Those which are expressive merely by custom or convention, for example the minor triad [indexical representation].

Some empirical evidence for this view has been provided in the study by Francés & Bruchon-Schweitzer (1983) which was referred to earlier (p.53). The results that they obtained were factor-analysed and suggested the following:

1. The feelings expressed by the pieces of music closely corresponded to the feelings visually expressed by the actors' bodies, regardless of the condition of induction (either verbal or musical). However, this similarity was limited to the activation dimension or kinetic aspect of emotional expression.

2. Distinguishing the affective tone (positive/negative) of expressed emotions which are similar in activation level depends on more conventional and culturally variable aspects of music which cannot be directly translated into body movement.

The suggestion here, then, is that different aspects of emotion (for example, hedonic tone versus activation level) may be conveyed by different kinds of musical cues. But how, then, do we explain differences of opinion - whether of one person with another, or of a person with his or her former self - about the emotions expressed by musical works? One possibility suggested by Sloboda (1985) is that the emotional cues in

music form a hierarchy, with primitive iconic cues such as tempo and volume available to most people regardless of their level of musical sophistication, and with more subtle (symbolic or indexical) cues such as aspects of tonal relations available only to those with deeper analytic powers with respect to the music. Therefore as one becomes more sophisticated musically (and perhaps obtains greater knowledge of the emotional world), so one acquires an improved ability to understand finer emotional nuances in music. Thus, while an inexperienced listener often finds the music of Mozart elegant but rather insipid and all vaguely 'happy' and trivial, especially when set beside the music of the Romantic composers, closer knowledge of Mozart's music typically results in an appreciation of its emotional richness, diversity and subtlety (Brown, 1980; Sloboda, 1985; Storr, 1992).

According to such cognitive psychological theories of music, while some of music's meaning may well be intrinsically specified by musical structure, much of the variance in the identification of meaning is attributable to characteristics of the listener and their familiarity with the musical idioms of their culture. Music is seen as a complex field of experience which draws upon cognitive, aesthetic, emotional and aural abilities (Peters, 1990), and which requires a *structured* listening. From this point of view, it is possible to fail to understand the emotional quality of a piece of music, since the ability to understand music's affective meaning depends on the requisite learning history as well as on the necessary reflective, analytic and interpretative abilities (Cunningham & Sterling, 1988).

What, then, are the variables related to the listener that are necessary to appreciate the emotional qualities of music?

## 3.3. The effects of listener characteristics on musical meaning

The effects of listener characteristics on the understanding of emotional expression in music have received little systematic attention. What information we do have, however, suggests that listener judgments are little affected by musical training

(Beldoch, 1964; Campbell, 1942; Edmonston, 1966; Hevner, 1935; Meerum Terwogt & van Grinsven, 1988, 1991), although it may be that in the randomly selected, small samples typically used in these studies within-group differences in training are not large enough to make their effects noticeable.

Gender, too, typically has no effect, or only a small and barely significant effect which is limited to particular emotions (Beldoch, 1964; Cunningham & Sterling, 1988; Dolgin & Adelson, 1990; Edmonston, 1966; Kastner & Crowder, 1990; Kratus, 1993; Meerum Terwogt & van Grinsven, 1988, 1991), while the effects of intelligence (operationally defined in a variety of ways) are inconsistent and unclear. Beldoch (1964), for example, found that the accuracy with which subjects identified the emotions expressed in music was related to their verbal intelligence as measured by a vocabulary test. Hevner (1935), on the other hand, found no relation between subjects' ability to distinguish between the major and minor modes in music and their scores on the Minnesota College Ability tests.

The ability to identify the broad emotional category into which a piece of music falls also does not seem to depend on familiarity with the composition (either having heard it before, or being able to identify it) (Hampton, 1945; Lieberman, Lewis & Walters, 1968). However, an interesting light has been thrown on the question of familiarity by Brown (1981), who designed a novel study aimed at getting away from a total reliance on verbal responses. Brown's subjects were asked to group nineteenth century orchestral and operatic excerpts into pairs that were maximally similar in meaning. The performance of experts and nonmusicians did not differ when the emotional characters of the excerpts were quite disparate. However, when listeners were asked, in a study entitled *Twelve Variations on Sadness*, to make subtle distinctions between musical excerpts with very similar emotional meanings and which fell into the same *verbal* emotion category, then above-chance agreement was only achieved by listeners who liked, and were well acquainted with, the musical styles and genres involved. A possible reason for this has been suggested by Sloboda (1985), who remarks that judgments of the emotional meaning of a piece of music are not necessarily made instantaneously as we understand a word or phrase in a spoken

language. Rather, the full appreciation of the emotional meaning of a piece of music might require careful and repeated hearings in the same way that repeated readings are necessary to appreciate the particular implications and connotations of a certain phrase in a poem.

Other listener variables which have been hypothesised to influence the interpretation of musical meaning but which have not yet been empirically tested include the existing mood of the listener, personality, motivation for listening to music, and attitudes toward music in general as well as towards the particular excerpt being listened to.

There is another approach, however, to exploring the effects of nature and nurture on the interpretation of emotional meaning in music. The focus of this approach is on the role of maturation and learning in the ability to understand one's community's music. Here the sort of questions that are asked concern the age at which the standard emotional connotations of music begin to appear, and whether these connotations become more reliable with age. These questions will be considered in the next chapter.

# CHAPTER 4:
# DEVELOPMENTAL TRENDS IN THE INTERPRETATION OF EMOTION IN MUSIC

## 4.1. Effects of age

As with all studies of musical meaning, some of the findings with children may reflect the particular musical stimuli used (recorded orchestral excerpts, solo instrumental pieces, songs, or specially composed melodies), as well as the response measures employed (usually verbal responses versus schematic drawings of facial expressions). Nevertheless, there is considerable agreement among studies that even preschool children can recognise a few basic emotions in music at better than chance levels. And by the time they are six or seven, children typically show beyond-chance agreement for all the emotions tested - most commonly happiness, sadness, anger and fear (Cunningham & Sterling, 1988; Dolgin & Adelson, 1990; Kastner & Crowder, 1990; Meerum Terwogt & van Grinsven, 1988, 1991; Trehub, Cohen & Guerriero, 1987 in Trehub, 1993). In addition, all the studies conducted to date except that of Kratus (1993) have found a developmental trend: the older children are, the more accurate they are at recognising the emotions expressed in music. Cunningham and Sterling (1988), Dolgin and Adelson (1990) and Meerum Terwogt and van Grinsven (1988, 1991) all found that children's ability to identify the emotions of happiness, sadness, anger and fear in music improved with age. Cunningham and Sterling (1988), for example, asked subjects aged 4, 5, 6, and 19-24 to assign verbal labels to orchestral excerpts previously judged by adults to express one of the above four emotions. The four-year-olds showed beyond-chance agreement with adults for only two of the emotions, the five-year-olds for three emotions, and the six-year-olds for all affective categories. Dolgin and Adelson (1990) investigated the ability of 4-, 7- and 9-year-old children to recognise the same four emotions in specially composed melodies presented in two modalities: soprano voice and viola. They gave their subjects the opportunity to express their choices either nonverbally (by selecting facial

expressions) or verbally. They found that four-year-olds' agreement with adults was above chance except for males' identification of fear played on the viola and females' identification of anger and fear played on the viola. Using a somewhat different approach, Scherer and Oshinsky (1977) conducted a pilot study in which they asked ten second-graders, nine sixth-graders and ten adults to choose between alternative labels (illustrated by stylised facial expressions) for twelve synthesised tone sequences that had been found consistently to receive particular emotional labels in their main study with adults. Their results showed a linear trend for age, with decreasing numbers of errors for older subjects.

## 4.2. Effects of emotion

As with adults, some emotions were found to be easier for children to identify than others. The general finding is that children at all age levels apparently find it easiest to identify happiness in music followed by sadness, anger and fear, usually in that order (Cunningham & Sterling, 1988; Dolgin & Adelson, 1990; Kastner & Crowder, 1990; Meerum Terwogt & van Grinsven, 1988, 1991; Scherer & Oshinsky, 1977). There is widespread agreement that children as young as four can identify happy music (Cunningham & Sterling, 1988; Dolgin & Adelson, 1990; Kastner & Crowder, 1990; Meerum Terwogt & van Grinsven, 1988, 1991; Trehub et al., 1987, in Trehub, 1993). However, children are less accurate at identifying other emotions in music. Cunningham and Sterling (1988) found that four-year-olds had difficulty recognising sadness in music, while both four- and five-year-olds were unable to recognise anger at above-chance levels. Cunningham and Sterling's five-year-olds and Dolgin and Adelson's four-year-olds also had difficulty identifying fear.

Although relatively few studies have examined children's errors, Meerum Terwogt and van Grinsven (1988, 1991) and Cunningham and Sterling (1988) found that anger and fear were frequently confused with each other. For example, in Cunningham and Sterling's (1988) study, one half of the four-year-olds' responses to the three anger segments were afraid interpretations, while only one-quarter were anger interpretations. In addition, Dolgin and Adelson (1990), using specially composed

melodies reflecting auditory characteristics associated with emotions in speech, found that their subjects tended to confuse frightened and sad-sounding melodies; furthermore, angry-sounding melodies were, if not correctly recognised, confused with happy-sounding ones, but not vice-versa. Dolgin and Adelson also pointed out that although the four-year-olds made many errors, their responses were not random. They avoided confusing emotions that according to theories such as those of Plutchik (1980, in Dolgin & Adelson, 1990) are 'polar opposites': happiness and sadness, and anger and fear. They concluded that four-year-olds generally understand some portion of the emotional content of music even when they cannot make a completely correct judgment. Taking a slightly different approach, Kratus (1993) found that children aged six to twelve were able to make more consistent distinctions between happy and sad music than between excited and calm music. Similarly, a table in Flowers' (1988) study (cited in Kratus, 1993) indicates that children in preschool through sixth grade who were asked to describe music used the adjectives 'happy' and 'sad' more frequently than they did 'exciting' and 'calm', although college students frequently used words such as 'excited', 'relaxing' and 'peaceful' to refer to music.

How can we explain these findings? If, as has been suggested above, musical meaning depends both on structural features of the music and on certain listener characteristics, then we would expect that an improvement in the ability to detect emotional expressiveness in music would be influenced firstly by an improved auditory processing capacity (in other words, a more advanced or differentiated perception of musical elements), and secondly by an increased knowledge of the meaning of emotion concepts in a particular community.

## 4.3. What do children know about music, and how does this knowledge develop?

It has recently become clear that the musical aspects of speech are particularly salient for infants, and that even newborns can distinguish between notes of different pitch and intensity (Fernald, 1989; Trainor & Trehub, 1992). By the time they are about 5 months old, infants can detect pitch changes as small as one semitone (the smallest ratio separating notes in Western tonal music), and can discriminate changes in both melodic contour and rhythm pattern (Chang & Trehub, 1977b in Trehub, Bull & Thorpe, 1984; Cohen, Thorpe & Trehub, 1987). By 19 months most children also produce songs with melodic contours and occasional, if irregular, rhythm patterns (Davidson, McKernon & Gardner, 1981).

There is evidence to suggest that children's initial grasp of music is 'topological': they categorise sequences of sound on the basis of global, relational properties such as melodic contour and rhythm, rather than on the basis of specific details (Trehub, 1987; Trehub, Bull & Thorpe, 1984; Trehub, Thorpe & Morrongiello, 1987). As children grow older, they gradually advance from using these global processing strategies to using more subtle features such as tonality and interval size. Although it is unclear when the ability to use these more subtle features appears, Davidson et al. (1981) found that by the age of five-and-a-half children could produce songs organised within a stable key or tonality. Five-year-olds can also distinguish closely-related keys from distant keys, although they struggle to detect changes in interval size in melodies (Bartlett & Dowling, 1980). At this age children also acquire the ability to perform expressive transformations, for example to slow down the pace of a song for a sad version thereof (Davidson et al., 1981).

On the basis of these findings, Gardner (1973a, in Hargreaves, 1986) argues that by the age of six or seven children possess many of the fundamental skills required for full-scale musical perception and performance. There is evidence that at least by the age of eight, children can not only distinguish between a melody played in different

keys, but can also detect interval changes in melodies (Bartlett & Dowling, 1980). It seems, therefore, that after the age of eight musical development involves merely an improvement in the basic skills that have already been acquired.

It is not very surprising, then, that children apparently base their judgments of the emotions expressed by music on much the same formal musical elements as those used by adults. When Meerum Terwogt and van Grinsven (1988, 1991) asked their subjects to justify their classifications of musical excerpts into emotion categories, they found that older subjects mentioned more and more varied musical elements such as tempo or the use of particular elements, but that all age groups mentioned rhythm and melody with equal frequency.

Kratus (1993) conducted a set of stepwise regression analyses to investigate the relationship between musical elements and emotion judgments, and found that schoolchildren listening to excerpts from Bach's Goldberg Variations based their happy-sad distinctions largely on the rhythmic activity (regardless of tempo) and articulation (staccato/legato) of the excerpt, while excited-calm distinctions were largely based on rhythmic activity and meter (duple/triple). Trehub et al. (1987, in Trehub, 1993) played very simple melodic sequences composed of either two or nine notes to children between the ages of 4 and 8 as well as to adults, and asked them to rotate a pointer to a sad face, a happy face, or a neutral face. For children, like adults, high pitch - fast rate two-note sequences and ascending nine-note sequences typically led to happy judgments, while the typical response to low-slow and descending sequences was sad.

The effects of modality on affective judgments by children are less clear. Trehub et al. (1987) found that children's responses to nine-note sequences using notes from either the major or minor scale were inconsistent. Kastner and Crowder, however, obtained contrasting results. They played twelve short musical passages in either the major or minor mode to 38 children between the ages of three and twelve, and asked them to point to one of four schematic faces chosen to symbolise happy, sad, angry and neutral/contented expressions. They found that the conventional positive

(happy/neutral) - major and negative (sad/angry) - minor connotation was present for all their subjects, even the youngest, and that the means differed significantly from chance. These results need to be interpreted with caution, however, as Felleman et al. (1983) have found a tendency for children to misidentify neutral expressions as sad. Nevertheless, some support for Kastner and Crowder's conclusions is provided by Kratus (1993). In his study, modality did not enter any of the stepwise regression analyses as a significant predictor of subjects' interpretations, as it was highly correlated with other variables that were more strongly related to subjects' responses. However, modality was significantly correlated with the percentages of both happy and excited responses, indicating that children associated the minor mode with both sad and calm responses. In this way, children interpreted modality in music in much the same way as adults do. Finally, Kastner and Crowder (1990) also found an interesting interaction between mode and type of arrangement (harmonised versus unharmonised). Minor tunes were perceived as sad or angry much more frequently when presented in their harmonised version than in their purely melodic version. And for major-mode stimuli, subjects associated unaccompanied items more often with positive emotions than accompanied items.

The mode of presentation also appears to have an effect: Dolgin and Adelson (1990) found that their youngest subjects, four-year-olds, were better able to recognise emotion in sung, rather than instrumental, melodies. It needs to be remembered, however, that Dolgin and Adelson's melodies were specially composed to include features associated with the expression of emotion in the human voice. It is therefore impossible to determine from this study alone whether this difference in accuracy should be attributed to the difference in the familiarity of voice and viola, whether it is intrinsically easier to interpret these particular emotional cues in voices than in musical instruments (or at least the viola), or whether both differences are important. Dolgin and Adelson's results also showed a significant condition × emotion interaction: sad melodies were more readily identified when performed on the viola, while angry melodies were more recognisable when sung.

Taken together, these findings provide some support for Gardner's (1973a) assertion that

> a reasonably competent 7-year-old should understand the basic metrical properties of his musical system and the appropriate scales, harmonies, cadences and groupings....What is lacking is fluency in motor skills, which will allow accurate performance, experience with the code, tradition and style of that culture, and a range of feeling life.
>
> (quoted in Hargreaves, 1986, p.83)

It is this "feeling life", or more specifically the child's knowledge and understanding of emotions, to which this discussion will now turn.

## 4.4. What do children know about emotions, and how does this knowledge develop?

Knowledge about emotions develops enormously between birth and adulthood, and the answer to the question of what children know about emotion therefore varies quite considerably depending on the age of the children in question. Furthermore, because emotions are multidimensional constructs it is difficult to make generalisations about the nature of emotional development as a unitary phenomenon, since each of the components of emotion may have different developmental timetables (Brody, 1985). As Manstead (1993) points out, however, research conducted within the last few decades suggests that at any given age children tend to know more about emotion than was previously thought.

Most developmental researchers have tended to focus on only one component of emotion such as children's emotional vocabulary, recognition of facial expressions, or knowledge of the causes and consequences of emotion. Data obtained by Bretherton, Fritz, Zahn-Waxler and Ridgeway (1986) and Smiley and Huttenlocher (1989) indicate that emotion language first emerges around eighteen months to two years of age, and increases during the third year. During this time children come to label the emotions of others as well as their own, to discuss past and future emotions,

and to talk appropriately about the antecedents and consequences of emotions. It has been suggested that at this stage children's emotion concepts merely refer to expressive behaviours, and that it is only later that they recognise that emotions also have internal, intrapsychic components (Beckwith, 1991; Smiley & Huttenlocher, 1989). However, Wellman et al. (1995), who examined children's use of emotion terms in natural language, have presented data to make a strong case that even children as young as two years appreciate that emotions are internal, subjective, experiential states, and distinguish them from the situations that elicit them and the expressions and actions that they give rise to.

Because many events that communicate affective information in early childhood are non-verbal and interpersonal, much of the research on children's understanding and awareness of emotions has focused on children's ability to label and recognise facial (and occasionally other non-verbal) expressions of emotion. Children have been given drawings, photographs or videos of posed or spontaneous facial expressions and been asked to provide or select a verbal label for them (Felleman, Barden, Carlson, Rosenberg & Masters, 1983; Field & Walden, 1982; Profyt & Whissell, 1991; Stifter & Fox, 1986; Tremblay, Kirouac & Dore, 1987), to select a facial expression that goes with a given label (Smiley & Huttenlocher, 1989), or to match facial expressions to a standard (Walden & Field, 1982). They have also been asked to select the emotions communicated by non-verbal vocal expressions (Dimitrovsky, 1964; McCluskey, Albas, Niemi, Cuevas & Ferrer, 1975; Stifter & Fox, 1986).

A second body of research has investigated children's ability to identify the emotions elicited by verbally described situations and to identify the causes and consequences of emotions. In this research tradition children have typically been asked to provide a verbal label, select a verbal label, or choose a photograph of a facial expression to represent the emotion that would be experienced by the protagonist in a short story (Borke, 1971, 1973; Camras, 1980; Camras & Allison, 1985; Gibbs & Woll, 1985), or to describe the causes and/or consequences of a set of emotions (Bortolotti, D'Elia & Whissell, 1993; Russell, 1990; Stein & Jewett, 1986).

Despite the methodological variability and focus on different components of emotion, however, the data indicate that there are several clear trends in the development of emotional understanding. The most consistent finding is that children's recognition and comprehension of emotional situations and expressions becomes increasingly sophisticated with age (Borke, 1971, 1973; Bortolotti et al., 1993; DePaulo & Rosenthal, 1978; Dimitrovsky, 1964; Felleman et al., 1983; Field & Walden, 1982; Gibbs & Woll, 1985; Manstead, Anderson & Bromley in Manstead, 1993; McCluskey et al., 1975; Profyt & Whissell, 1991; Russell, 1990; Smiley & Huttenlocher, 1989; Tremblay, Kirouac & Dore, 1987; Walden & Field, 1982; for reviews see Gross & Ballif, 1991 and Manstead & Edwards, 1992).

Several studies suggest that pre-verbal infants can discriminate between positive and negative facial expressions and that they can associate positive and negative facial expressions with vocal expressions of emotion, although one cannot be sure that specific emotion meaning (e.g. 'fear' as opposed to merely 'something negative') is understood at this stage (Field, Woodson, Greenberg & Cohen, 1981 in Field & Walden, 1982; LaBarbera, Izard, Vietze & Parisi, 1976; Walker, 1982; Young-Browne, Rosenfeld & Horowitz, 1977). By about two-and-a-half years of age, the majority of children can identify happy facial expressions, but have difficulty distinguishing among other emotions. By the age of about four or five children can generally recognise happy, sad, angry, scared and surprised facial expressions at above-chance levels, although they still perform relatively poorly compared to adults. By this time they can also identify situations that are likely to cause these emotions in others at above-chance levels, and can produce largely plausible - by adult standards - antecedents and consequences for a variety of positive and negative emotions. The ability to distinguish between specific facial expressions within the general categories of positive or negative expressions continues to improve with age, and by the time they are nine years old, children typically display a very high level of accurate judgments. Although much less is known about children's abilities to interpret others' auditory emotion signals, there is evidence from studies starting at age five or six that recognition ability for vocally expressed emotions also improves with age at least until age 10 or thereabouts, when the rate of improvement declines slightly

(Dimitrovsky, 1974; McCluskey, Albas, Niemi, Cuevas & Ferrer, 1975; Stifter & Fox, 1986). Older children also perform with greater accuracy than younger children when given a combination of audio and visual channels of nonverbal communication (DePaulo & Rosenthal, 1978, 1979).

A second conclusion is that, across ages, children identify certain emotions correctly more often than others. Although there are some differences among studies, the general result is that happiness is differentiated, identified and labelled earlier and more consistently than other states, usually followed by sadness and anger, and subsequently by fear, disgust and surprise. Emotion terms requiring more complex understanding of interpersonal situations (for instance, gratitude, guilt or pride) are not appropriately used until the age of about seven or eight (Borke, 1971, 1973; Bortolotti et al., 1993; Camras & Allison, 1985; Felleman et al., 1983; Field & Walden, 1982; Gibbs & Woll, 1985; Harris et al., 1987, in Manstead, 1993; Profyt & Whissell, 1991; Smiley & Huttenlocher, 1989; Stifter & Fox, 1986; Tremblay, Kirouac & Dore, 1987; Walden & Field, 1982).

Although relatively few studies have analysed children's errors, it appears that their misjudgments are systematic. Although children can make accurate discriminations between positive and negative emotions, there is a general tendency for them to confuse emotions within these broad categories. For example, Borke (1971, 1973) and Felleman et al. (1983) found a tendency for sadness to be misidentified as anger and vice versa. Similarly, Manstead et al. (in Manstead, 1993) found that when children were asked to match situations with emotion words or facial expressions, their youngest subjects responded to fear and anger with sad labels, and to disgust with sad or angry judgments. Manstead et al. (1993) and Russell (1990) also found a trend for situations which the majority of subjects linked with excitement to be described as leading to happiness.

Other studies have found errors to be characterised by similarity along the dimension of activity, rather than valence. Dimitrovsky (1964), for example, found that vocal expressions of active emotions (anger and happiness) tended to be confused, as did the

passive emotions of sadness and love. Felleman et al. (1983) found a tendency for children to confuse sad and neutral facial expressions, and a secondary tendency for anger to be misidentified as happiness.

In other cases, both the activation and the pleasure dimensions seem to be involved. Green and Ekman (1973, in Smiley & Huttenlocher, 1989) found that four-year-olds had some difficulty distinguishing sad from angry, angry from scared, and surprised from both scared and happy facial expressions when asked to select an expression that matched a given situation. The last finding parallels that of Manstead et al. (in Manstead, 1993), who point out that it reflects the ambivalent nature of the concept 'surprise' on the dimension of pleasure-displeasure.

There is also some suggestion that children's emotion concepts are not necessarily equivalent to those of adults. In particular, children may interpret surprise, and perhaps also excitement, more positively than do adults. For example, according to the Facial Action Coding System (FACS) developed by Ekman and Friesen (1971), Field and Walden's (1982) photograph of a child's expression for surprise was a happy/surprise blend. Russell (1990) and Bullock and Russell (1984) have found that children are much more likely than adults to give a positive sense to the word 'surprise', although they treat a facial expression of surprise as neither positive nor negative. Similarly, Stein and Jewett (1984) reported the results of a pilot study showing that young children associate excitement and surprise with positive events.

Felleman et al. (1983) point out, however, that an important precaution in interpreting results, which is absent in most studies of children's ability to recognise emotional states, concerns analysis of the role played by basic inclinations to use, or not to use, various emotional labels. They found that children had high base rates for using the labels 'happy' and 'sad', which increased their accuracy in recognising those states, while low base rates for anger and neutral labels hindered accuracy in identifying those states. Similarly, Manstead et al. (in Manstead, 1993) found that younger children used the terms happy and sad significantly more often than older children, perhaps explaining why, although age trends were seen for other emotions, younger

children were as accurate in identifying happiness and sadness as the older children. In another study, Dimitrovsky (1964) found that children identifying vocal expressions of emotion responded more often with the negative emotions (sad and angry) than with the positive emotions (happy and loving), which partly accounted for their greater accuracy in recognising the negative states.

Several other methodological considerations also need to be taken into account when evaluating the above findings. Different results across studies may be due to the fact that the various techniques used for measuring children's recognition and understanding of emotion differ considerably in the task demands placed on subjects. Several researchers have expressed concern that children's performance on emotion recognition tasks may be confounded by limitations in their verbal ability: higher recognition scores are usually found on research tasks that require the comprehension of emotion words than on tasks that involve explanations of and/or the production of emotion words (Gross & Ballif, 1991; Lewis, 1989). There is also evidence, however, that specifying emotions using either only words or only facial expressions may produce different responses, and that at least for some emotions such as fear and disgust, verbal labels are in fact understood better that facial expressions (Camras & Allison, 1985; Manstead et al. in Manstead, 1993; Russell, 1990).

In those studies that do use facial expressions, the quality of the stimulus materials may also be an issue. Children may not respond in the same way to line drawings, static photographs, and videotapes (see, for example, Field & Walden, 1982), and static representations of facial expressions may have little relation to the subtly changing facial expressions that occur during everyday interactive situations, so the external validity of these measures is questionable. This problem is exacerbated by the fact that some researchers do not evaluate their stimulus materials to make sure that the intended emotion is in fact portrayed.

A similar problem occurs when children are asked to link emotions with situations. It has been pointed out that the stories presented to children may be ambiguous, or that they might conceivably lead to more than one emotional reaction. For example, a

scenario in which a playmate breaks the child's toy is typically expected to lead to anger (at the playmate's destructive behaviour); however, it is equally plausible that it could lead to sadness (because the toy is broken). One's emotional reaction will depend on which aspect of the situation one focuses on, possibly reflecting individual differences in responding to frustration (Borke, 1971, 1973; Camras, 1986; Field & Walden, 1982). Michalson and Lewis (1985, in Lewis, 1989), for instance, presented their subjects with a situation of a child lost in a store, assuming that the 'correct' emotional response would be fear. However, their adult subjects were equally divided as to whether this situation would lead to fear or sadness, and their child subjects associated the situation primarily with sadness: these responses could not really be described as 'incorrect' simply because they disagreed with the researchers' interpretations.

Another methodological limitation is that many studies do not compare children's response patterns with those of adults, making it difficult to determine when the developmental process has reached its peak (Lewis, 1989). Developmental advances during later childhood may be obscured if we assume that adults' responses would be the same as those of older children in the study. Alternatively, not knowing adults' responses might lead us to underestimate children's emotion knowledge. For example, if we find that only sixty percent of five-year-old children label a facial expression correctly, we could easily assume that the developmental process is incomplete at this age if we did not know that adults also only get the correct solution sixty percent of the time.

In spite of these methodological limitations, however, a few firm conclusions can be drawn. It appears that the ability to identify specific emotions in both music and in other, more 'natural' stimuli emerges in the preschool years and improves with age. Accuracy does, however, vary with the particular emotion depicted, and inaccuracies or misidentifications tend to be systematic rather than random.

## 4.5. Explaining children's emotional development

### 4.5.1. Traditional approaches

How do we account for these findings? Unfortunately, although the kinds of studies discussed above have added much to our knowledge of children's understanding of emotions, few researchers have looked at how children come to know what they know. The majority of studies are not grounded in or motivated by a particular theory of emotional development, and most researchers therefore briefly attempt to explain their findings about the changes that occur in children's knowledge of emotional behaviour on the basis of post hoc theories. This often leads to inconsistent data and confusion, making it difficult to compare studies and systematically accumulate knowledge. Without a theoretical rationale, it is unclear how we can explain the mechanisms or processes that might underlie the way in which children's interpretation of a particular emotional situation, facial expression, or piece of music develops.

Those explanations that have been provided for the way in which emotional knowledge changes with age have typically grown out of the theorist's explicit or implicit beliefs about the nature of emotion. For example, the psychoevolutionary or biosocial theorists' assumption that there is a set of discrete, universal, biologically-based emotional expressions each of which involves distinctive types of facial, vocal, gestural and bodily expression (Izard, 1991) has led scholars who support this position to claim that human infants are endowed with innate and instinctive recognition devices for interpreting the meaning of this repertoire of emotional expressions (Darwin, 1904). This line of thinking suggests that children's early concepts of emotion are probably anchored to facial - and (although we know less about them) perhaps other nonverbal - expressions. Thus, irrespective of culture, the first emotions a child will come to understand will be 'basic' emotions such as happiness, sadness and anger that can be conveyed easily by facial expressions (Harris & Saarni, 1989). During development, children's understanding of emotion becomes more

sophisticated as they associate these nonverbal expressions with distinct feelings, and modify them by associating them with particular behaviours and symbols (including verbal labels) as they mature and take part in an increasing variety of social experiences (Malatesta & Izard, 1984 in Gross & Ballif, 1991). According to this view, there is a tight and inherent link between a particular situation and a particular emotion. The role of the social context, therefore, is to facilitate or inhibit the emergence of universal cognitive abilities rather than to be an important determinant of emotional development in its own right (Gordon, 1989; Harris & Saarni, 1989). Because research in this tradition has tended to assume that such words as 'anger' and 'fear' refer to a set of schemata that are biologically grounded and effectively culture-free, an understanding of emotion is typically seen as a straightforward registration of a universal truth. As a result, certain responses are considered correct, and the rest errors. The focus of research within this tradition has therefore tended to be on how accurately and rapidly children come to recognise the universal and biologically-based links between different emotion components such as expressions and experiences.

If we apply this position to music, it might be postulated that there are innate, universal auditory 'emotion-detectors' common to both the prosodic aspects of speech and music (e.g. Scherer, 1979; Scherer & Oshinsky, 1977; Trainor & Trehub, 1992), just as there are innate emotion detectors for facial expressions. Izard (1991), for example, argues that each basic emotion involves distinctive changes in tone of voice (and specifically in the fundamental frequency, or pitch thereof) or possess other individual acoustic characteristics. Perhaps, then, children's errors in labelling emotional pieces of music can be attributed to children's perceptual limitations regarding their ability to recognise and differentiate each of the basic emotions as expressed in the auditory modality.

This explanation for children's errors has been applied to the visual modality by Stein and Jewett (1986), who argue that children may experience difficulty in labelling facial expressions because they have difficulty locating and encoding the unique features of each basic emotional expression which adults supposedly can detect. Therefore children might have more difficulty identifying complex facial expressions

like fear (which involves seven independent muscle movements) than simple expressions like happiness, which involves only two muscle movements (Matsumoto, 1992). In addition, they might systematically confuse expressions in which the positions of facial features are similar (for example, anger and disgust).

Some evidence for this idea has been obtained by Walden and Field (1982). In their research on children's understanding of facial expressions, they found that pointing out facial features to children aided discrimination of expressions that they found problematic. In addition, De Paulo and Rosenthal (1978, 1979) have found evidence that non-verbal decoding skills become more differentiated with age, with older subjects being able to process, integrate and utilise more information than younger subjects.

In contrast to the biological theorists, the social constructionists argue that the links between the components of emotion (for example, facial expressions and situations) are socially constructed and therefore arbitrary: no particular component is essential to the whole. In other words, terms such as anger imply not an absolute truth, but rather a culturally relative one. Consequently, the more radical proponents of this position claim that the child can have little understanding of emotion that is not socially transmitted, as society determines what emotions will be experienced, what their subjective experiences will be like, and how they will be interpreted. Therefore children's emotional knowledge changes not only because of cognitive development, but also because they construct their own emotion knowledge in response to their changing social interactions, their exposure to the beliefs, vocabulary, norms, metaphors, rules and so on provided by their culture, as well as through the effects of reinforcement and observational learning (Gordon, 1989; Harris & Saarni, 1989).

According to this view, children's understanding and awareness of particular emotions is primarily the result of socialisation: children learn about certain emotions through observing their own and others' behaviour, and through verbal explanations by others which may either reveal or conceal certain norms and beliefs about those emotions (Harris & Olthof, 1982 in Gordon, 1989). Although social constructionists have

studied variations in emotional experience and expression linked to social structure and child-rearing practices, their findings have not been linked to the development of children's understanding of emotion. However, Borke (1971, 1973) suggested quite persuasively, on the basis of her findings regarding American and Chinese children's abilities to label emotional situations, that emotion recognition is influenced by the salience and social acceptability of particular emotions in different societies. This position implies that children would have more difficulty recognising certain emotions - whether in situations, facial expressions, or music - that are less acceptable to them, or that have low social acceptability in their culture.

This explanation has an advantage over the biological theories in that it can explain why children show similar patterns of performance across different expressive modalities as well as when labelling emotional situations. In fact, virtually all theories of emotional development do make at least some provision for the influence of social learning, and many theorists would agree with Borke's view despite thinking the social constructionist view to be too extreme. What this argument cannot explain, however, is why certain emotions are systematically confused with each other. Thus while there seems to be some truth in both the biological and the social constructionist perspectives, neither alone seems able to provide a full account of how and why children's knowledge of emotions develops.

### 4.5.2. An integrative approach

The next logical step would therefore appear to be to try and integrate these two approaches, which is what James Russell and his colleagues have done (Bullock & Russell, 1984, 1986; Russell & Bullock, 1985, 1986; Russell & Ridgeway, 1983; Russell, 1989). According to their position, which has been referred to as 'moderate' social constructionism, children do not possess innate, discrete emotion categories. Nevertheless, there are universal constraints on their interpretation of emotion. Young children, it is argued, initially classify emotions not in terms of discrete categories but rather in terms of two continuous dimensions: degree of pleasure and (either simultaneously or slightly later) degree of arousal or activation. In other words,

young children may initially interpret emotion expressions in a global, relatively undifferentiated way according to their general tone, rather than in terms of adultlike categories labelled anger, fear, and so on. This hypothesis does not deny that very young children discriminate and categorise different emotions, but argues that children may initially discriminate between emotion expressions in quantitative rather than qualitative terms.

With development, then, children have to learn to differentiate within their global interpretation of the emotional domain until they acquire the adult meanings of the categories labelled anger, fear, and so on. The child is thought to develop from interpreting emotions in terms of pleasure and arousal to distinguishing emotions located at similar points in the two-dimensional space through acquiring emotion *scripts*. Cognitive gains and interpersonal experiences (including observation and direct teaching) allow children to learn that the features that make up emotion events (for example, contexts, causes, beliefs, physiological reactions, desires, subjective feelings, facial expressions, behaviours and consequences) are correlated and tend to occur in predictable sequences. In addition, they are stimulated to differentiate among emotions previously treated as if they were alike by the labels and interpretations their society provides. Thus, while the ability to interpret emotions in terms of broad dimensions is argued to be innate and pancultural, children learn how to classify and differentiate among emotions from their particular culture.

This view of development has two implications for children's understanding of emotion concepts. Firstly, although emotion categories are hypothesised to be fuzzy at all ages, the boundaries of emotion categories become less fuzzy with increasing age. Younger children have broader conceptions of emotion categories than their older counterparts, and initially overgeneralize emotion words to all emotion states similar in pleasure and arousal (which explains why they make more 'errors' when labelling facial expressions). As they grow older, children's conceptions of what define an emotion become narrower and more specific. Secondly, because emotion scripts are based on the existence of characteristic or frequently recurring features as opposed to necessary and sufficient universal features, the points within the two-

dimensional grid at which children develop emotion prototypes will depend on their culture and may change with development as their emotional experiences change.

According to this view, children's 'errors' in recognising and labelling emotional expressions are manifestations of their conceptual system: expressions similar in pleasure and arousal may initially be treated as alike. For example, what appears to an adult as fear, anger, tension, disgust or upset may all appear to the child as 'unpleasant and aroused' (Russell & Bullock, 1985, 1986). Consequently, these authors argue that at any given age, some categories such as happiness are used more consistently because they overlap less with other categories presented to the child - 'happiness' is usually the only positive emotion included in studies and hence is maximally distinct from the other, less pleasant, emotions.

This theory argues that children will begin to distinguish between states similar in pleasure and arousal as they acquire the adult emotion scripts for these states. Some simple scripts, such as 'happiness' (which involves a pleasant feeling due to attaining a desired state) will be acquired before those of more complex emotions. Very young children may therefore treat terms such as anger and sadness (both of which involve goal failure and negative feelings) as though they have the same meaning because they have not yet acquired the key components of the 'adult' emotion scripts, which according to Stein and Levine (1989, in Manstead, 1993) involve a focus either on the causes of the goal failure (leading to anger) or on the consequences thereof (leading to sadness). Similarly, older pre-schoolers might not make an adult-like differentiation between words such as surprise, excitement and happiness, presumably because they often hear the first two words in pleasant contexts (for example, "I've got a surprise [gift] for you"). It is only later that they learn that words such as surprise and excitement refer to high arousal states which are not necessarily pleasant. School-aged children and adults usually have little trouble making this distinction, but may treat emotions such as guilt and shame - which have even more complex scripts requiring a differentiation on the basis of internal versus external causes of the displeasure at having violated a rule - as though they had the same meaning. Their failure to make fine distinctions between closely related emotions does not, however,

mean that such distinctions are not possible. Rather, from this perspective the potential taxonomy of emotion concepts is infinite.

Russell and his colleagues have provided intriguing evidence that when children are asked to select a photograph of a facial expression that depicts a particular emotion, their errors are quite systematic. They interpret this as support for their contention that even before they have any knowledge of emotion names or can label facial expressions of emotion 'accurately', very young children nevertheless interpret emotions in a meaningful way and can categorise and judge the similarity between them. In an unpublished study (reported in Russell, 1989) in which 2-year-olds were asked to associate photographs of facial expressions with words like happy or angry, these children distinguished positive from negative facial expressions, but were unable to distinguish among the negative expressions. Similar findings have been obtained by other researchers (Borke, 1971; Camras & Allison, 1985; Felleman et al., 1983; Manstead et al. in Manstead, 1993; Stifter & Fox, 1986; Smiley & Huttenlocher, 1989).

Additional evidence that the pleasure and arousal dimensions contribute to children's understanding of emotion comes from factor analyses and multidimensional scaling of children's self-reported emotions (Russell & Ridgeway, 1983) and from multidimensional scaling of the perceived similarity ratings between different facial expressions. In a set of studies utilising the latter statistical technique, Bullock and Russell (1984, 1986; Russell & Bullock 1985, 1986; Russell & Ridgeway, 1983) found that the errors made by children between the ages of two and four (as well as by adults) were quite systematic and conformed to what the authors predicted on the basis of their circumplex model of emotions, in that errors were likely to be a direct function of the distance between the target emotion and the other emotions in terms of pleasantness and arousal. In other words, the errors made in facial expression tasks were more likely to be selections of 'neighbouring' emotion categories in the structural model than other category choices, with the probability of a choice decreasing with distance (that is, the number of steps) away from the target emotion. Supporting evidence that pleasure and arousal dimensions underlie even young

children's emotion concepts has been provided by Whissell and D'Elia (1993) and Moland and Whissell (1993) who found that adults' open-ended categorisations of children's drawings of, and statements of antecedents of, six emotions could be predicted on the basis of ratings of pleasantness and arousal. These researchers also found that the distance between emotions in two-dimensional space was predictive of confusion between different emotions.

Russell and his colleagues have also found evidence that for most emotion words, children's categories narrow with age. For example, in a study reported by Bullock and Russell (1986) children were shown photographs of facial expressions and asked "Is this person x?" where x was one of 9 emotion terms. On average, a facial expression was put into 4.0 categories by three-year-olds, 2.5 categories by four-year-olds, and 2.2 categories by five-year-olds. When children were asked to select the best expression to match a given label, younger children also tended to be overinclusive by adult standards - that is, they overgeneralized emotion words to all emotional facial expressions similar in pleasure and arousal. For each emotion there was a focal point, the expression most likely to be chosen, with the probability of a choice decreasing with distance from the focal point. However, the range around the focal point was broader for younger than for older subjects, and became narrower with increasing age. This finding is in accordance with the research of Alexander and Enns (1988) who (although they were not concerned with emotions) found similar effects when they examined the boundaries of categories formed by preschoolers and adults following exposure to a continuum of new objects (stuffed animals). Although category boundaries were fuzzy at all ages, they became relatively less fuzzy (that is, increasingly exclusive) with increasing age. Younger children overextended categories to a broader range of exemplars than older subjects, even if this meant that the two categories were not mutually exclusive, and underextended concepts to similar exemplars.

Russell and Bullock's theoretical model leaves important questions about emotional development unanswered. Firstly, as Manstead (1993) points out, it is not clear why and how the young child's concept of emotion is constrained by the pleasantness and

arousal dimensions, or whether the ability to make distinctions in terms of these dimensions is innate or learned. Secondly, there is a lack of clarity about how scripts and dimensions are related, as well as about what emotion scripts need to contain in order for the child to make adult-like discriminations among emotion concepts. The research reported by Harris (1989) indicates that children's representations of emotion go beyond associations that they have built up from what they have experienced directly or vicariously to include some knowledge of the role played by their own and other's beliefs and desires. As Manstead (1993) argues, this suggests that change in the ability to attain or maintain a valued state might play a larger role in emotion scripts than Russell and Bullock's model allows. Stein and Levine (1989) and Wierzbicka (1992a, 1992b), for example, have argued that emotions can be defined primarily in terms of what people want and don't want (that is, their goals and values). In addition, of course, Russell and his colleagues' analysis only concerns one aspect of emotional development: that of emotion concepts or categories, particularly as they relate to facial expressions. As these authors themselves note (Russell & Bullock, 1984), a full account of how children come to interpret emotional expressions requires knowledge of a number of other aspects of emotion. These might include topics such as the ways in which children interpret and reconcile conflicting emotional information from different sources (for example, facial expressions and situations), the specific facial, vocal, or gestural features influencing the perception of emotion, the relationship between felt and expressed emotion, the effects of children's upbringing and experiences on their emotion knowledge, and so on.

Without denying these limitations, however, it does seem that Russell and his colleagues' model of the development of children's conceptualisation of emotions is consistent with much of the empirical evidence on this topic briefly reviewed above, and that this model provides a useful starting point for future research. This is not to deny that there may well be other explanations which do not fit clearly into any of the theoretical frameworks discussed above which may help to account for children's interpretation of emotion in music as well as in other modalities. For example, some researchers have argued that the feelings children attribute to music, to story situations, or to facial expressions may partially reflect their own affective

experiences. If children are cognitively egocentric and incapable of distinguishing self from other, they may be incapable of differentiating their own emotional response (such as fear to an angry facial expression or piece of music) from the emotion expressed in the stimulus itself (Brody, 1985; Meerum Terwogt & van Grinsven, 1988, 1991). Since the relationship between perceived and felt emotion appears to be particularly complex in relation to music, even for adults, it is therefore unclear whether children who are asked to respond to music are reporting how the musical sequences made them feel, or whether they understand, in adult-like fashion, that such patterns simply convey emotional meaning.

A major advantage of adopting Russell and his colleagues' approach to studying children's understanding of emotions, however, is that it encourages psychologists to question, and if necessary discard, the implicit assumptions that have dominated thinking and research regarding children's understanding of emotion. In particular, these researchers argue that investigating what children do, accurately or not, when asked to interpret emotional expressions, and searching for any order in the processes or strategies that may underlie their responses, is likely to be more revealing than simply evaluating children's recognition of emotional expressions as correct or incorrect against an adult standard (Bullock & Russell, 1986). These authors point out that although accuracy scores can give us some information, they may not tell us everything. Consequently, rather than making the assumption that children are doing exactly the same thing as adults do, but simply making more mistakes, we must be careful to interpret their responses in terms of their own emotion concepts, not those imposed by adults. If the boundaries between emotion concepts are fuzzy, for example, a child may meaningfully associate more than one emotional label with a particular piece of music. There need be no inherent, one-to-one correspondence between emotion words and musical excerpts, although we would expect children's conception of what musical excerpts express a given emotion to become narrower, or more exclusive, with age. Additionally, if emotion categories are interrelated according to an underlying structure, children who provide 'inaccurate' labels might still be interpreting expressive music in an orderly fashion related to the underlying dimensions of degree of pleasure and activation. From this perspective, the important

question is whether children's responses when they are asked to label, categorise, or discriminate among emotions in music are systematic. If children's responses are systematic, then it becomes possible to specify the features that underlie their interpretations. We can examine their emotion concepts - the ways in which they categorise, organise, and interpret emotional experience - and how these concepts develop.

## Conclusions

The relationship between music and emotion is notoriously complex and raises some perplexing questions for psychologists. Although research in this area has produced some interesting findings, many questions remain unanswered - partly because of the conceptual and methodological difficulties facing researchers interested in these issues, but partly also because the majority of studies have lacked an explanatory theoretical framework. If we want to understand how the ability to interpret emotions in music develops over the life-span, we need to develop suitable theories which can be empirically tested.

One theory of emotional development which has the potential to contribute to a greater understanding of the cognitive and affective systems involved in coming to understand music's emotional meaning is that developed by Russell (1980, 1991b) and his colleagues. Their theory, incorporating the idea that emotion concepts are prototypically organised and fuzzy, as well as interrelated according to their position on two or more underlying dimensions, promises to provide fertile ground for exploring the ways in which people think about and use emotion concepts to describe music, and hence to increase our understanding of both the psychology of music and emotional development in general.

# CHAPTER 5: RATIONALE AND METHOD

## 5.1. Rationale

The goal of the present study was to further our understanding of both emotional development and people's understanding of emotion in music by extending Russell and his colleagues' research on children's understanding of emotional facial expressions (Bullock & Russell, 1984, 1986; Russell, 1989; Russell & Bullock, 1985, 1986; Russell & Ridgeway, 1983) to the more culturally variable stimulus of music. In particular, it aimed to examine whether a theoretical framework incorporating the ideas of the fuzziness and structural organisation of emotion concepts is useful in explaining how children come to apply emotion terms to music, and how their interpretation of emotion in music changes as they grow older.

Underlying the choice of this approach was the assumption that because the emotional 'meaning' of musical excerpts is ambiguous, in categorising them subjects would be required to make decisions about which, and how many, particular features a musical excerpt needs to possess in order to be classified as expressing a particular emotion, and which features are dispensable or irrelevant. It was felt that a study which goes beyond a simple investigation of the effects of increasing age on 'accuracy' to investigate the *nature* of subjects' responses and what underlies them, should be able to tell us something about how children and adults conceptualise emotion categories, as well as to increase our understanding of what it is that people are actually doing when they assign emotion labels to music.

Specifically, the following research questions and hypotheses were posed:

**Research question 1: At what age does the ability to interpret emotions in music emerge?**

Because young children's identification of emotion in music appears to be influenced by methodological variations and the particular task demands placed on them (for example, the number and nature of emotion categories used as response alternatives, response measures [verbal or nonverbal] employed, particular musical excerpts used and medium of performance [e.g. vocal versus instrumental]), it was difficult to formulate a specific hypothesis regarding the age at which children would become able to identify particular emotions in the musical selections to be used in this study. Nevertheless, on the basis of previous research findings (Cunningham & Sterling, 1988; Dolgin & Adelson, 1990; Meerum Terwogt & Van Grinsven, 1988, 1991) it was expected that even preschool children would be able to identify at least some emotions in the music at above-chance levels.

**Research question 2: Do younger and older children and adults differ in the consistency with which they interpret emotions in music?**

Based on a theoretical prediction that fuzzy category boundaries will become less fuzzy (that is, will become more specific, narrower, and socially agreed-upon) with increasing age (Alexander & Enns, 1988; Bullock & Russell, 1984, 1986; Russell & Bullock, 1985, 1986), and extrapolating from the fairly consistent findings of previous research (Cunningham & Sterling, 1988; Dolgin & Adelson, 1990; Meerum Terwogt & van Grinsven, 1988, 1991), the following hypothesis was proposed:

**Hypothesis 1: Consistency among different individuals in the emotional categorisation of musical excerpts will increase with age.**

Previous studies have, however, consistently found that agreement among individuals regarding the emotions expressed by different musical excerpts depends not only on the age of the subjects, but also on the particular emotion being expressed (Campbell, 1942; Cunningham & Sterling, 1988; Dolgin & Adelson, 1990; Kastner & Crowder, 1990; Meerum Terwogt & van Grinsven, 1988, 1991; Scherer & Oshinsky, 1977). This led to the third research question:

**Research question 3: Do children and adults interpret some emotions in music more consistently than they interpret others?**

Based on the findings of the above studies, the following hypothesis was formulated:

**Hypothesis 2: Subjects will show greater agreement for the categorisation of 'happy' excerpts than for excerpts expressing other emotions.**

A study by Felleman et al. (1983) investigating children's recognition of emotional facial expressions suggests that differential accuracy for particular emotions may reflect adults' and children's varying tendencies to use particular emotion labels more frequently than others, rather than true discriminative ability. The following question was therefore posed:

**Research question 4: Do basic inclinations to use, or not to use, various categories of emotional labels affect the degree of agreement among children and adults?**

Although previous studies have not examined the question of *intra*-individual consistency, fuzzy category theory also predicts that the boundaries of emotion concepts will become more stable and specific for each individual as they grow older - a prediction that has been supported by research investigating children's categorization of a set of novel stuffed animals (Alexander & Enns, 1988). Consequently, the following hypothesis was proposed:

**Hypothesis 3: Older subjects will show greater individual consistency than younger subjects in the emotion labels assigned to particular musical excerpts across different categorisation tasks.**

Although the consistency of judgments was expected to increase with age and to be greater for particular emotions, the theoretical framework on which this study is based leads to the prediction that even adults will not be perfectly consistent in their judgments, and that the degree of agreement will vary for different musical excerpts. This is because if subjects' categories are organised around prototypes rather than classically defined, category membership should be graded - that is, we would expect

different musical excerpts to be better or poorer musical exemplars of a certain emotion category. More specifically, the following prediction was offered:

**Hypothesis 4: Inter-subject consensus will be greater for some musical excerpts than others, with some excerpts being fairly prototypical members of a particular emotion category and others being more peripheral or borderline members.**

Furthermore, Bullock and Russell's work on children's understanding of emotional facial expressions (1984, 1986; Russell & Bullock, 1985, 1986) suggests that in some cases children's ideas about what constitutes a prototypical member of a particular emotion category might change as they grow older. This possibility was considered in the next research question:

**Research question 5: Will the emotion category into which the majority of subjects place a particular musical excerpt change systematically with age?**

Not only does the theoretical framework for this study predict that subjects' prototypes may change with age, however, but it also predicts that with increasing age, children's categories will become narrower and more specific - in other words, less fuzzy. Consequently, we would expect that the range of emotions around the modal choice selected by subjects will narrow with age. More specifically, the following proposal was made:

**Hypothesis 5: With increasing age, the number of emotion categories selected by subjects as applicable to describe a given musical excerpt will decrease.**

A prototype analysis of emotion concepts predicts that because emotion concepts overlap and the boundaries between them are fuzzy, even when children and adults are 'inaccurate' in their judgments - that is, when their judgments do not agree with the modal choice - they are nevertheless interpreting the emotional qualities of music in an orderly and predictable way. In other words, we would expect their 'errors' to be systematic rather than random. More specifically, Russell and Bullock's (1985, 1986;

Bullock & Russell, 1984, 1986) circumplex model predicts that confusion between different emotions will be a function of their similarity on the underlying bipolar dimensions of degree of pleasure and degree of activation.

In practical terms, what this means is that emotion labels similar in degree of arousal and pleasure will all be more-or-less applicable to a given musical excerpt, with the probability of a choice decreasing as one moves further away on these two underlying dimensions (Bullock & Russell, 1984, 1986; Russell, 1991b; Russell & Bullock, 1985, 1986).

This led to the following prediction:

**Hypothesis 6: For each musical excerpt there will be a focal point, the emotion category most likely to be chosen, with the majority of errors falling into the two adjacent emotion categories, and probability of a choice decreasing with distance (in terms of pleasure and activation coordinates) from the focal emotion.**

In addition, if the dimensions of pleasure and activation/arousal do underlie the emotion concepts of subjects of all ages, we would predict that they will be able to make reliable distinctions between musical excerpts expressing similar degrees of pleasure and different degrees of activation, as well as between those expressing similar degrees of activation but different degrees of pleasure. Furthermore, the emotion categories into which they place each excerpt should reflect the degree of pleasure and arousal expressed by that excerpt. This led to the formulation of the next two hypotheses:

**Hypothesis 7: Subjects of all ages will be able to make reliable distinctions between musical excerpts varying in degree of either pleasantness or activation, when the other underlying dimension is held more-or-less constant.**

**Hypothesis 8: Subjects' emotional judgments will be correlated with independent ratings of the degree of pleasure and arousal expressed by each musical excerpt.**

Assuming that subjects can make reliable distinctions between musical excerpts in emotional terms, the next question that arose concerns the basis for these judgments.

**Research question 6: On what musical and/or extra-musical features do subjects of different ages base their judgments of the emotions expressed in musical excerpts?**

Here the focus was on whether the same features are used by children and adults, or whether children restrict themselves to fewer features, weigh the same features differently, or concentrate on different features. A related question concerns whether some emotion 'cues' in music are more natural or unlearned than others. Although theoretical disagreements and inconsistencies in empirical findings did not permit the formulation of specific hypotheses, it was expected that there would be a relationship between the number of subjects associating each excerpt on the stimulus tape with particular emotions and the formal musical characteristics of the excerpts.

Finally, since previous research has shown marked individual differences in nonverbal decoding skills (Davitz, 1964a; Meerum Terwogt & van Grinsven, 1991) the present study aimed to make an exploratory attempt to find correlates of individual subjects' ability to interpret emotions in music. Specifically, it was decided to investigate the following characteristics:

**Research question 7: Is there any relationship between the ability to interpet emotions in music and subjects'**
**(a) gender**
**(b) intelligence or academic achievement**
**(c) ability to interpret expressions of emotion in speech, and**
**(d) musical training?**

The research design that was developed in an attempt to answer these questions and test the hypotheses proposed is outlined in the following section.

## 5.2. Method

### 5.2.1. Subjects

The sample consisted of 120 subjects divided equally among four age levels: 5-year-olds (12 girls and 18 boys, mean age = 5 years 6 months), 7-year-olds (16 girls and 14 boys, mean age = 7 years 6 months), 9-year-olds (13 girls and 17 boys, mean age = 9 years 6 months), and adults (23 women, 7 men, mean age = 33 yrs 5 months, range = 18 - 46). The children were drawn from a preschool and primary school in a predominantly lower-middle class suburb of Cape Town, both of which catered to a fairly broad range of socio-economic classes and races. The 7- and 9-year-olds were randomly selected from Sub-A (Grade1) and Standard 1 (Grade 3) classes at the primary school. The 5-year-olds and adults were volunteers, the former recruited verbally from two preschool classes and the latter recruited through written advertisements at suburban libraries, on the university Psychology Department notice boards, and at a children's ballet studio.

One 7-year-old girl failed to complete tasks 2 and 3 as she was absent from school on the second day of testing. Seven 5-year-olds failed to complete all the tasks for the same reason, but were replaced by other children of the same age and sex.

### 5.2.2. Stimuli

Music Tape

The experimenter selected an initial pool of 48 music excerpts, each 25-30 seconds long, chosen to represent a wide range of musical styles, genres, and traditions and subjectively judged to vary in emotional character. These included music drawn from children's and adults' movies, musical theatre and ballets, classical 'art' music, contemporary 'pop' music (including local African music), electronic music, jazz, ragtime, Scottish and Spanish folk music, and melodies specifically composed for toddlers. Dates of composition ranged from the early eighteenth century to the 1990s.

Instrumental versions of all pieces were used to avoid confounding due to the meaning of lyrics. The excerpts were recorded on a Marantz stereo sound system and played to subjects on a portable radio cassette recorder, Sharp WQ-283Z.

A pilot group of 5 adults with varying degrees of musical experience listened to the excerpts and provided open-ended descriptions of their emotional character. Eight excerpts which these pilot subjects did not perceive as being emotionally expressive, or which evoked widely differing descriptions, were excluded from the study and replaced with alternative excerpts selected by the author.

A second group of 15 musically knowledgeable adults then rated this revised set of 48 excerpts on Mehrabian & Russell's (1974) semantic differential scales for arousal and pleasure, which are presented along with the instructions given to subjects in Appendix 1. In order to avoid subject fatigue, each pilot subject rated only a third of the total set of musical excerpts; each excerpt was therefore independently rated by five individuals. For each of six combinations of dimensions (high arousal-pleasure; high arousal-displeasure; moderate arousal-pleasure; moderate arousal-displeasure; low arousal -pleasure; low arousal-displeasure) dimensional ratings were examined in the light of the emotional descriptions provided by the initial 5 pilot subjects in order to select three excerpts: two clear or good examples of each category, and one less clear or borderline example. This technique was used partly to avoid bias in the selection of musical exemplars to fit the author's preconceived notions of their emotional character, and partly to ensure that the musical excerpts did indeed express varying degrees of pleasure and activation. In addition, any excerpts that evoked a very negative reaction from the pilot subjects, or whose title or source (e.g. a particular movie soundtrack) had been correctly identified by one or more of the pilot subjects were excluded from the final selection. This procedure was followed in an effort to reduce confounding due to familiarity with, and/or specific associations to, the music.

The final stimulus selection therefore consisted of eighteen musical excerpts of 25-30 seconds each, separated by 10-second intervals of silence. This set of excerpts was

recorded three times (once for each of the tasks described on pp. 99-100 below). On the first tape, the 18 excerpts were divided into two sets of 9 excerpts expressing similar degrees of pleasure/displeasure. On the second tape, they were divided into three groups of 6 excerpts expressing similar degrees of activation. On the third tape, the entire set of musical excerpts was presented as a whole. The order of the excerpts was randomised on each tape. Table II summarises the title, composer and/or performer and album title of the excerpts, their corresponding mean arousal and hedonic tone ratings, and the order of presentation for Task 3. It also provides illustrative examples of the emotional descriptions of these excerpts provided by the pilot subjects.

The musical excerpts were independently analysed by the author and 3 music students to provide a description of the formal musical elements in each piece. The analytical constructs used were similar to those previously investigated. The following characteristics were analysed on 5-point scales:

TEMPO: 1 = largo, 5 = presto

RHYTHM: 1 = flowing, 5 = firm

RHYTHMIC ACTIVITY: 1 = little, 5 = much

ARTICULATION: 1 = legato, 5 = staccato

PITCH LEVEL: 1 = low, 5 = high

PITCH RANGE: 1 = small, 5 = large

HARMONY: 1 = consonant, 5 = dissonant

The judges also rated the excerpts for modality (major, minor or atonal).

Interrater reliability for the 5-point scales was calculated by correlating the ratings from all possible combinations of the four judges. The mean correlation coefficients obtained were as follows: Tempo: $r = 0.86$; Rhythm (firm vs. flowing): $r = 0.81$; Rhythmic Activity: $r = 0.82$; Articulation: $r = 0.60$; Pitch Level: $r = 0.21$; Pitch Range: $r = 0.47$; Harmony: $r = 0.64$. These coefficients were significantly greater than would be expected by chance with alpha limited to 0.01 for all elements except Pitch level and Pitch range, which were subsequently eliminated from further analyses.

**Table II:** Composition details, mean hedonic tone and arousal ratings and illustrative descriptions by pilot subjects of musical excerpts

| | Title | Composer | Date | Album Title | Hedonic tone | Arousal | Pilot subjects' descriptio |
|---|---|---|---|---|---|---|---|
| 1. | The Fight | S. Prokofiev | 1934 | Romeo and Juliet | 3.2 | 15.6 | agitation. exuberanc |
| 2. | Wild Bears | E. Elgar | 1908 | The Wand of Youth Suite, No. 2, Op. 1b | 5.6 | 13.0 | happy excitemen panic |
| 3. | Desolation Theme | Leigh Harline | 1940 | Walt Disney's *Pinocchio* | -5.2 | -5.4 | thoughtful ness, nostal; |
| 4. | Nocturne No. 8 in D flat major | F. Chopin | | | 11.0 | -8.8 | peacefulne: sadness |
| 5. | Hero's Wedding | Patrick Doyle | 1993 | William Shakespeare's *Much Ado About Nothing* | 12.0 | 10.0 | triumph, gra happiness |
| 6. | Have I told you lately | Van Morrison | | The Best of Van Morrison | 11.8 | -7.2 | |
| 7. | Dance of the four knights, from *Checkmate* | Arthur Bliss | | Three English Ballet Suites | 0.8 | 6.8 | menace, arousal |
| 8. | Journey to the Sun | D. Hewitt | 1989 | An African Tapestry | -2.0 | -8.8 | tranquility calmness, boredom |
| 9. | Shake! | Mango Groove | | Another Country | 10.8 | 11.0 | |
| 10 | The Whale Chase | Leigh Harline & Ed Plumb | 1940 | Walt Disney's *Pinocchio* | -4.2 | 13.8 | fear, confusion |
| 11 | Romance | S. Prokofiev | 1933 | Lieutenant Kije: Symphonic Suite Op. 60 | -2.2 | -6.6 | melancholy nostalgia |
| 12 | Who wants to live forever? | Brian May | 1986 | Queen: Greatest Hits II | -0.8 | -1.0 | |
| 13 | Ragtime Dance | Scott Joplin | | The Sting | 1.4 | 6.2 | clumsy happiness jollity |
| 14 | The fall of Saigon, from *Miss Saigon* | Alain Boublil & Claude-Michel Schönberg | | The Royal Philharmonic Orchestra play suites from *Les Miserables & Miss Saigon* | -5.8 | 13.8 | dread, dominance pomposity |
| 15 | La Cabane de Baba - Yaga sur des Pattes de Poule | Modest Mussorgsky | 1874 | Pictures at an Exhibition | -5.8 | 6.2 | agitation, expectanc |
| 16 | El Relicario | (Traditional) | | The Soul of Spain - 101 Strings | 7.6 | 12.8 | |
| 17 | The Ugly Sisters | Paul Reade | 1993 | Cinderella | 0.4 | 15.2 | excitemen humour |
| 18 | Abraham's Theme | Vangelis | 1981 | Chariots of Fire | 4.4 | -9.8 | tranquility expectanc |

Reliability was considered acceptable for the remaining 5 elements. Interrater reliability for modality was established by calculating the percentage of agreement for all possible combinations of the four judges. Mean interrater reliability was 80%, which was also considered acceptable.

Means of each element for the eighteen excerpts were computed. For modality, excerpts were coded as 1 if at least three of the four judges rated them as minor, and as 2 if at least three of the judges rated them as major. No modality coding was given to five excerpts in which modality could not be clearly established or which were rated by two or more judges as atonal.

Control Tape

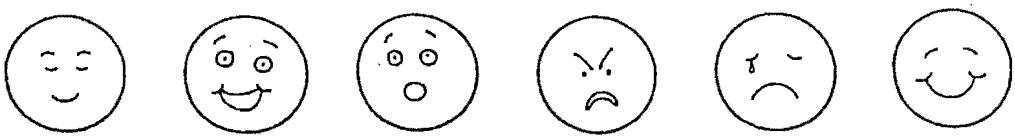
A tape consisting of six segments of verbal/vocal expressions of emotion was constructed for purposes of comparison. A female actress read 6 short passages adapted by the author from novels with the intention of expressing each of the 6 emotions used in the study. These passages are presented in Appendix 2. As with the music tape, segments were separated by ten seconds of silence and presented in random order.

### 5.2.3. Response Measures

Because previous research (Camras & Allison, 1985; Russell, 1990) has suggested that children are able to identify emotions most accurately when presented with a combination of verbal labels and facial expressions rather than either alone, subjects of all ages were provided with a combination of schematic line drawings of emotional faces (shown in Figure 2) and the verbal labels of the emotion categories they were intended to represent. The six emotion labels used in the study were calm, happy, excited, scared, angry/cross and sad. These emotions were selected because they have all been successfully used in at least one previous study with children and represent a broad spread on the dimensions of arousal and pleasure, incorporating three positive and three negative emotions at each of three levels of arousal or activation (low,

moderate, and high). In addition, Ridgeway, Waters and Kuczaj (1985), who collected normative data on (English-speaking) children's ability to understand emotion-descriptive adjectives, found that the majority of 4-year-old children were reported by their parents to understand the above 6 terms: understanding ranged from 73.3% for 'calm' to 100% for 'happy', 'sad' and 'angry'.

**Figure 2:** Schematic facial expressions used as response alternatives



calm excited scared angry sad happy

In order to validate that the line drawings did express the intended emotions and that subjects could discriminate between them, 38 pilot subjects (20 females, 18 males) between the ages of 5 and 52 (mean = 17 years, 10 months) made forced choice ratings of either which of the six emotions was best expressed by each face, or which of the six facial expressions best expressed each emotion. Interrater reliability (percentage agreement) for each face/emotion pair was as follows: calm: 92.1%; happy: 89.5%; excited: 81.6%; scared: 94.7%; angry: 100%; sad: 100%.

### 5.2.4. Procedure

The general procedure used in the study involved asking the subjects to select the emotion word or facial expression that best matched the emotion expressed in each musical excerpt (or, for the control task, each verbal expression). In addition to completing the control task, each of the 120 subjects who participated in the study listened to the musical excerpts three times. On each of these three occasions of listening to the music, however, they were presented with different combinations of emotion categories from which to choose as response alternatives. On one occasion (labelled Task 1), they were asked to select one of three emotions expressing similar

degrees of pleasure but different degrees of activation. On another occasion (Task 2) they were asked to choose between two emotions expressing similar degrees of activation but different degrees of pleasure. Only on Task 3 were they presented with all six emotion terms simultaneously. Details of these tasks are provided below.

CONTROL TASK

Subjects' ability to understand the requirements of the tasks and to identify emotions expressed in speech was investigated by playing them the six verbal/vocal expressions of emotion which comprised the control tape, and asking them to select one of the six emotion word/face pairs described above that best matched the way the woman on the tape was feeling in each extract.

MUSIC TASKS:

TASK 1: Activation Dimension Varied

In order to determine whether children could use the dimension of level of arousal or activation to categorise musical stimuli, they were given three face-word pairs representing pleasant emotions varying in degree of arousal (that is, excited, happy and calm). They then listened to the 9 musical excerpts rated by the pilot subjects as expressing pleasant emotions, and were asked to select the facial expression or emotion that best matched the feeling in the music. This procedure was repeated for the nine musical excerpts previously rated as expressing negative emotions, with the subjects being asked to choose between scared, angry and sad expressions. Half the subjects began with the pleasant emotions, and half began with the unpleasant emotions.

TASK 2: Pleasure Dimension Varied

In order to determine whether subjects could discriminate between musical excerpts expressing similar degrees of arousal or activation but different degrees of pleasure,

they were asked to match the 6 musical excerpts which fell into the high arousal category with either excited or scared facial expressions, to match the 6 excerpts which fell into the moderate arousal category with either happy or angry facial expressions, and to match the six excerpts which fell into the low arousal category with either calm or sad facial expressions.

TASK 3: Both Dimensions Varied

Subjects were presented with all six emotional expressions used in the study. After hearing each of the eighteen musical excerpts, they were asked to select the face-word pair which best matched the feeling in the music.

The order in which the control task and the three music tasks were completed was randomised among subjects.

The actual procedure followed in carrying out these tasks was slightly different for the five-year-olds and the older subjects. The five-year-olds were tested individually. At the beginning of each trial, the experimenter shuffled cards containing the emotional facial expressions and spread them out in a random array on the table in front of the child. Before playing the stimulus tape in the first session, the experimenter pointed to each of the cards and named the emotion depicted. At each of the subsequent sessions the experimenter made sure that the child could remember what emotion each face was intended to represent. Children were reminded of any emotion labels for the facial expressions that they had not recalled spontaneously. After hearing each excerpt on the tape, the child was asked to point to the face that best matched the feeling in the music and the experimenter recorded the child's picture choice. The stimulus tape was stopped if necessary between each excerpt to give the child time to respond. Response times varied from almost immediate (before the excerpt was completely finished) to several seconds after the music had ended. All responses were praised.

Older children and adults were tested in small groups of up to eight persons, and were instructed to circle their choices on an answer sheet which (depending on the task) consisted of 6, 9 or 18 boxes containing the names of either 2, 3 or 6 emotions as well as the line drawings of faces depicting each of the emotions. In an effort to minimise the effect of response order on the results, the order of facial expressions was different in each box. The directions given to subjects were based on the model in Appendix 3, with slight modifications according to the nature of the specific experimental task and the age of the subjects.

In order to guard against subject fatigue and maintain the children's interest, preschoolers were tested in three separate sessions on different days, each lasting approximately 20 minutes. The older children were tested on two different days in sessions lasting approximately half-an-hour. All of the children were tested in a quiet room at their schools. Adult subjects were tested in a single session lasting approximately one hour, either in a room in the Psychology Department or at the author's home.

Additional measures

Because IQ scores were only available for a minority of the subjects, teachers' ratings of children's reading and mathematics performance (A, B, C or D) were obtained from the 7- and 9-year-old children's end-of-year report cards (which were produced in the same week as the testing) as an indicator of their intellectual achievement.

Adults were asked to indicate in writing the number of years formal musical training (defined as lessons on an instrument, singing lessons and/or music theory lessons) that they had had, as well as describing their participation in musical activities such as singing in a choir or playing a musical instrument. Information on the children's musical training was provided by the primary school's music teacher.

Qualitative Data

Although time restrictions and the limitations of younger children's vocabulary made questioning each subject about the reasons for each of their judgments impractical, it was felt that individuals' subjective impressions and experiences of the tasks could provide an interesting complement to the quantitative data and perhaps give rise to hypotheses for future research. Consequently, each of the experimental sessions with the adult groups was followed by a short, semi-structured discussion session in which subjects were asked for their general impressions of, and comments on, the tasks, as well as being asked specific questions concerning the techniques they used to make their judgments, any particular difficulties they experienced, and familiarity of the musical excerpts. Answers were recorded by the author in writing. In addition, any spontaneous comments made by the children during the experimental sessions which related to the above issues were recorded by the experimenter.

### 5.2.5. Scoring procedures

Because the emotional character of musical excerpts can seldom if ever be determined a priori and objectively, there have been two main approaches to judging the 'accuracy' with which subjects assign pieces of music to emotion categories. The first method is to select or specially compose musical excerpts which are subjectively judged by the experimenter or music experts to represent a particular emotion, sometimes also validating these judgments using forced-choice ratings by adults. The second approach is to avoid making prior judgments about the emotional meaning of musical excerpts in favour of seeking consensus among subjects, typically accepting the group modal choice as the 'correct' classification. Although both of these approaches have been useful, a consensus approach was adopted in this study for several reasons. One of these is that this research was concerned with how ordinary people interpret music's emotional character and whether this interpretation is orderly or meaningful, and not with whether this interpretation is 'accurate' in terms of expert opinion. Moreover, a composer's intended emotional meaning can seldom be objectively determined, but rather only inferred, and even experts do not always agree

with each other about the emotional character of musical excerpts. It was also felt that taking a consensus approach to scoring the results would reduce the likelihood of the experimenter's own preconceived notions of the emotional character of each excerpt affecting subjects' responses through demand characteristics.

Consensus can be measured in several different ways. As far as *intersubject* consensus (group agreement) is concerned, we can examine the frequency with which children in each age group labelled each musical excerpt in accordance with the modal emotion choice of their own age group, of all the subjects collapsed across age, or of the adult group who - as a result of their more advanced cognitive abilities and wider experiences of music and emotion, and so on - are assumed to be a more knowledgeable or accurate comparison group whose responses can be used as a standard. In addition to these potential measures of intersubject consensus, however, Alexander & Enns (1988) point out that a measure of *intrasubject* consensus (individual consistency) across different categorisation tasks is also potentially important when studying fuzzy categories, as it can indicate whether a fuzzy category boundary identified in a group analysis reflects disagreement among subjects who are themselves consistent in different contexts, or whether each individual lacks clear boundaries for the concept or category.

Consequently, it was decided to score each individual's responses in three different ways. In the first case, responses to the musical excerpts were scored as correct if they were in agreement with the adult modal choice for that excerpt. Because developmental trends and differences between the responses of adults and children of different ages were a particular focus of this study, this 'accuracy' score (Score A) was considered to be the variable of primary interest, and unless otherwise specified was used as the dependent variable in the statistical analyses described in the following chapter. However, two alternative scoring procedures were also used for comparison purposes. In one of these (Score B), children's responses were scored as correct if they were in agreement with the modal choice of their own age group rather than that of adults. In the other , those subjects who completed all three tasks (N = 112) were given an individual inconsistency score (Score C) representing the mean

number of emotional categories into which they placed each of the eighteen excerpts across the three tasks. Responses to the control task were scored as correct if they agreed with the emotion that each verbal segment was intended (by the author and the actress) to express.

# CHAPTER 6: RESULTS

## 6.1. Control task results

Scores obtained on the control task suggested that even the youngest subjects understood the nature of the tasks and had the requisite skills for solving them. A one-way analysis of variance of control task scores by age group found that there was a significant main effect of age ( $F(3, 116) = 16.183$, $p < 0.000$). Post-hoc comparisons conducted using Tukey's Honest Significant Difference (HSD) test with alpha limited to 0.05 indicated that 5-year-olds were significantly less accurate than the 7- and 9-year-olds and adults (who did not differ among themselves). Nevertheless, when one-sample *t*-tests were used within age levels to compare mean correct responses on the control task, presented in Table III, with that which would be expected by chance (M = 1), the means were significantly better than would be expected by guessing or random responding for all age groups including the preschoolers (all *p*s < 0.001).

**Table III.** Mean correct interpretations of control stimuli as a function of subject age and stimulus emotion (n = 30 for each age group)

| Subject Age (years) | Stimulus Emotion | | | | | | |
|---|---|---|---|---|---|---|---|
| | Calm | Happy | Excited | Scared | Angry | Sad | Total (max. = 6) |
| 5 | 0.53 | 0.40 | 0.30 | 0.40 | 0.73 | 0.27 | 2.43 |
| 7 | 0.50 | 0.60 | 0.77 | 0.37 | 0.93 | 0.83 | 3.67 |
| 9 | 0.80 | 0.77 | 0.80 | 0.90 | 1.00 | 0.83 | 4.47 |
| Adult | 0.77 | 0.47 | 0.93 | 1.00 | 1.00 | 0.83 | 4.03 |
| Combined | 0.65 | 0.56 | 0.70 | 0.67 | 0.92 | 0.69 | 3.65 |

## 6.2. Intersubject consistency

Traditionally, data on children's interpretation of emotion in music as well as in other contexts (for example, describing situations or labelling facial expressions) has been evaluated in terms of their agreement with an adult norm. It was therefore decided to begin with this approach and to investigate age differences in the number of subjects agreeing with the adult modal choice regarding the emotional categorisations of the musical excerpts (Score A).

### 6.2.1. Comparison with chance

Research question 1 asked at what age subjects' identification of emotion in music would begin to exceed chance performance. Mean correct responses to the music excerpts (defined as agreement with the adult norm, or score A) for each task and each age group are presented in Table IV. As was done with the control task, these were analysed to investigate whether they differed from what would be expected by chance.

**Table IV.** Mean correct interpretations of music stimuli for each task as a function of subject age

| Task | Response options | Maximum possible | Chance | Subject Age (years) | | | |
|---|---|---|---|---|---|---|---|
| | | | | 5 | 7 | 9 | Adult |
| 1 | Calm, Happy, Excited | 9 | 3 | 4.30 | 5.77 | 7.30 | 7.90 |
| | Sad, Angry, Scared | 9 | 3 | 5.17 | 6.17 | 7.43 | 7.87 |
| | Combined | 18 | 6 | 9.47 | 11.93 | 14.73 | 15.76 |
| 2 | Calm, Sad | 6 | 3 | 3.37 | 4.59 | 4.50 | 4.10 |
| | Happy, Angry | 6 | 3 | 4.13 | 4.62 | 5.10 | 5.57 |
| | Excited, Scared | 6 | 3 | 3.80 | 4.31 | 5.03 | 5.27 |
| | Combined | 18 | 9 | 11.30 | 13.51 | 14.63 | 15.40 |
| 3 | All 6 emotions | 18 | 3 | 6.93 | 8.21 | 11.07 | 12.57 |

When subjects' mean 'accuracy' score on each of the three tasks was compared with a score that would be expected with random responding (33.3% for task 1, 50% for task 2 and 16.67% for task 3), means of all age groups were found to be significantly greater than chance performance with alpha limited to 0.01. Division of tasks 1 and 2 into their component sub-tasks, however, indicated a single exception to this finding: 5-year-olds failed to perform at above-chance levels when asked to distinguish between sad and calm excerpts in task 2.

### 6.2.2. Comparison with adult mode

Research question two was concerned with age-related differences in agreement for the emotional categorisation of musical excerpts. Hypothesis 1 asserted that consistency among individuals in the emotion label chosen for the musical excerpts would increase with age. When consistency was defined as agreement with the adult norm (Score A), an initial inspection of the data suggested that this was in fact the case. Mean consensus among adults for the 18 excerpts averaged across the three tasks was 82%, and ranged from 43.3% to 100% for particular musical excerpts. The percentage of subjects agreeing with the adult modal choice was somewhat lower for children, ranging from 3.3% to 100% for nine-year-olds (M = 75%), from 6.9% to 100% for seven-year-olds (M = 62%) and from 10% to 96.7% for five-year-olds (M = 51%). The distribution of emotion choices for each extract, in each task, for each age group is presented in Appendix 4.

ANOVA Results

In order to investigate these effects in more detail, the accuracy scores obtained on the music tasks (Score A) were analysed using a 2-way analysis of variance for each task in which age was a between-subjects factor and emotion was treated as a within-subjects factor. In each of these three ANOVAs, age consisted of four levels: 5-year-olds, 7-year-olds, 9-year-olds and adults. Emotion consisted of six levels: calm, happy, excited, scared, angry and sad. The dependent variable, or 'accuracy' score, was the mean proportion of responses in agreement with the modal emotion choice of

the adult group for each emotion category.[1] These accuracy scores are presented in Table V, and are displayed graphically in Figure 3. The 0.05 level of probability was used to evaluate the main and interaction effects for the ANOVA. Follow-up analytic comparisons of means were computed using Tukey's Honest Significant Difference (HSD) test, with $\alpha$ also set at 0.05.

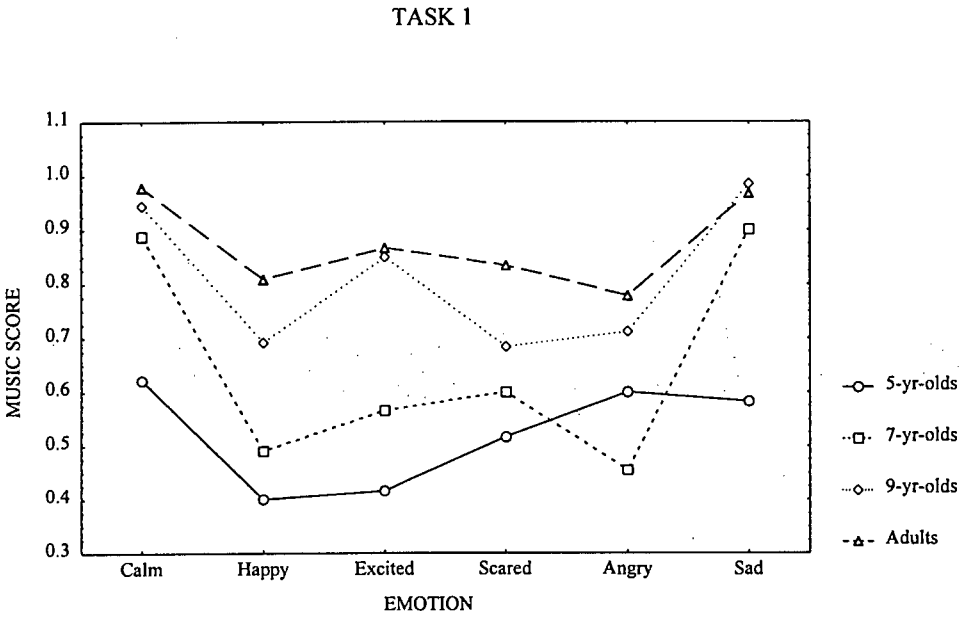
**Table V.** Mean correct interpretations of music stimuli (agreement with adult mode, expressed as percentages) as a function of subject age and stimulus emotion

| Task | Subject age (years) | No. of subjects | Stimulus emotion | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Calm | Happy | Excited | Scared | Angry | Sad |
| 1 | 5 | 30 | 62 | 40 | 42 | 52 | 60 | 58 |
| | 7 | 30 | 89 | 49 | 57 | 60 | 46 | 90 |
| | 9 | 30 | 94 | 69 | 85 | 68 | 71 | 98 |
| | Adult | 30 | 98 | 81 | 87 | 83 | 78 | 97 |
| | Combined | 120 | 86 | 60 | 68 | 66 | 64 | 86 |
| 2 | 5 | 30 | 54 | 65 | 60 | 70 | 67 | 54 |
| | 7 | 29 | 64 | 79 | 70 | 76 | 74 | 68 |
| | 9 | 30 | 70 | 76 | 82 | 88 | 95 | 68 |
| | Adult | 30 | 77 | 88 | 86 | 92 | 97 | 69 |
| | Combined | 119 | 66 | 77 | 74 | 81 | 83 | 65 |
| 3 | 5 | 30 | 38 | 32 | 25 | 63 | 38 | 58 |
| | 7 | 29 | 52 | 31 | 47 | 66 | 45 | 55 |
| | 9 | 30 | 66 | 57 | 48 | 93 | 51 | 73 |
| | Adult | 30 | 68 | 69 | 63 | 93 | 63 | 80 |
| | Combined | 119 | 56 | 47 | 46 | 79 | 49 | 67 |

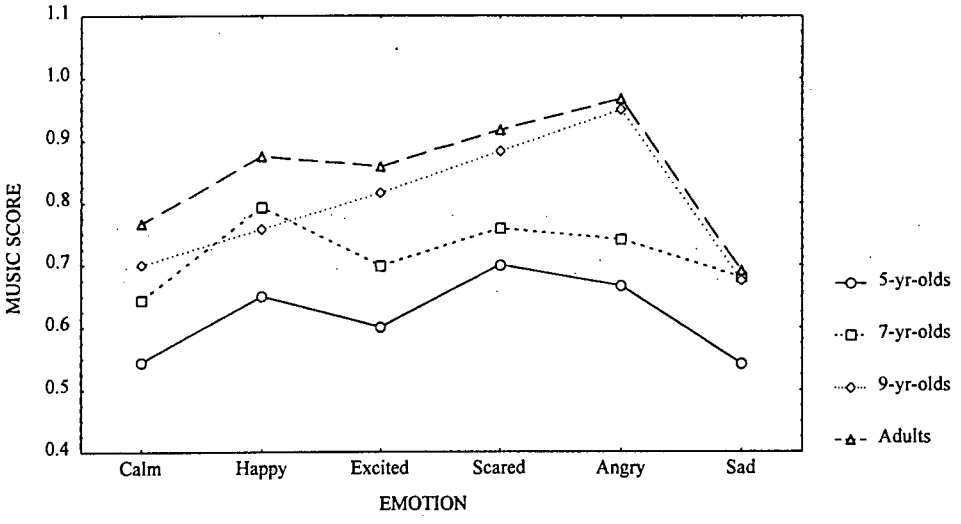
[1] Scores were expressed as proportions so that scores for different emotion categories would be comparable, as there were unequal numbers of musical excerpts in each emotion category according to the adult modal judgment.

The summary tables for the ANOVAs are provided in Appendix 5. For Task 1 (Activation dimension varied) significant main effects of Age ( $F(3, 116) = 36.429$, $p < 0.000$) and Emotion ( $F(5, 580) = 27.753$, $p < 0.000$) were qualified by a significant two-way interaction between Age and Emotion ( $F(15, 580) = 3.465$, $p < 0.000$). There were no significant interactions on Task 2 (Pleasure dimension varied) and Task 3 (Both dimensions varied). However, significant main effects were obtained for both Age (Task 2: $F(3, 115) = 34.350, p < 0.000$; Task 3: $F(3, 115) = 28.040, p < 0.000$) and Emotion (Task 2: $F(5, 575) = 18.386, p < 0.000$; Task 3: $F(5, 575) = 23.696, p < 0.000$) on each of these tasks.

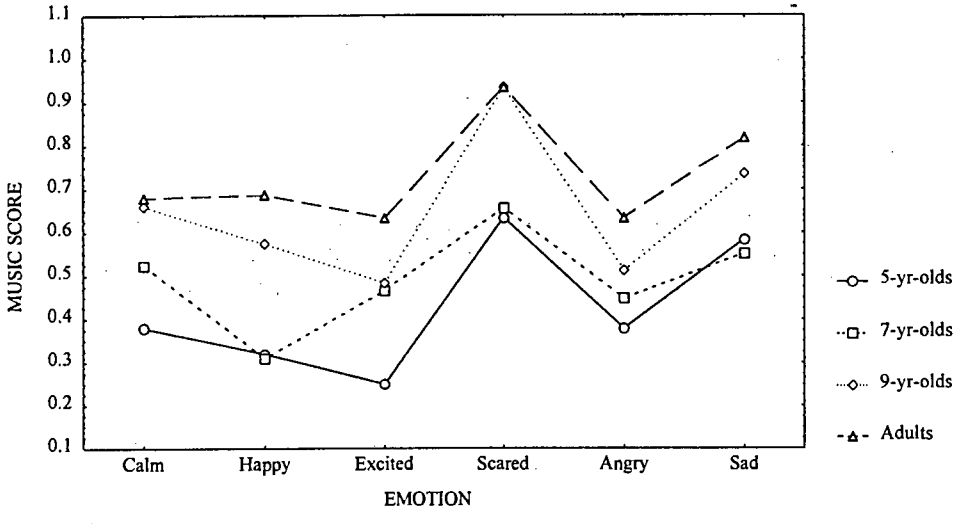
**Figure 3.** Plot of mean correct interpretations of music stimuli (agreement with adult mode) as a function of subject age and stimulus emotion



TASK 2




TASK 3




## Effects of Age

The plots of means presented in Figure 3 suggest that there was a general trend (despite a few exceptions) for 'accuracy' - that is, agreement with the adult norm - to improve with age, since the means of the older groups are generally higher than those

of the younger groups. However, the ANOVA results indicated that the effects of age were slightly different for the different tasks.

On task 1, analysis of simple main effects for age indicated that the effects of age were significant for all six emotion categories (all $ps < 0.01$). However, the effects of age on the number of subjects agreeing with the adult norm varied depending on the emotion being expressed. Post-hoc analytic comparisons of means (Tukey's HSD, $\alpha = 0.05$) indicated that for calm excerpts, 7-year-olds, 9-year-olds and adults were all significantly more accurate than 5-year-olds, but did not differ among themselves in their mean accuracy. Adults and 9-year-olds were significantly more accurate than 5-year-olds in identifying happy excerpts, and the adults were also more accurate than the 7-year-olds. Excited excerpts were identified correctly significantly more often by the adults and 9-year-olds than the 5- and 7-year-olds. Adults (but not 9-year-olds) were also more accurate than the 5- and 7-year-olds at identifying scared excerpts. Angry excerpts were identified correctly significantly more often by the adults and 9-year-olds than by the 7-year-olds, while adults, 9-year-olds and 7-year-olds were all significantly more accurate than the 5-year-olds in identifying sad excerpts.

The main effect for age obtained on task 2 was explored using analytic comparisons of means (Tukey's HSD test) with alpha set at 0.05. The results of these comparisons indicated that there were no significant differences between the mean accuracy scores of adults and 9-year-olds. However, adults were significantly more accurate than the 7-year-olds, who in turn were significantly more accurate than the 5-year-olds.

Post-hoc comparisons of means indicated that adults also did not perform with significantly greater accuracy than the 9-year-olds on Task 3. Both adults and 9-year-olds, however, were significantly more accurate than the 5- and 7-year-olds on this task (all $ps < 0.05$).

Effects of emotion

Research question 3 was concerned with whether the accuracy of subjects would vary depending on the emotions expressed by the musical excerpts. The analysis of variance results suggested that this was in fact the case, although the effects of emotion appeared to vary according to the task. However, no support was obtained for hypothesis 2, which predicted that happy excerpts would be identified with greater consistency than excerpts expressing other emotions.

On Task 1, the effects of emotion varied according to the age of the subjects. The performance of 5-year-olds and adults did not vary significantly for different emotions. 7-year-olds, however, were significantly more accurate at identifying calm and sad excerpts than any of the other four emotion categories (all $ps < 0.001$), while 9-year-olds identified the sad and calm excerpts correctly significantly more often than the happy, scared and angry excerpts (all $ps < 0.05$).

On Task 2, post-hoc comparisons of means used to explore the main effect of emotion indicated that calm and sad excerpts were identified correctly significantly *less* often than excerpts expressing all four other emotions (all $ps < 0.05$). In addition, angry excerpts were identified correctly significantly more often than excited excerpts ($p < 0.01$).

On Task 3, analytic comparisons of means indicated that the scared excerpt was identified correctly significantly more often than excerpts expressing all five other emotions, while the sad excerpts were also recognised correctly significantly more often than happy, excited and angry excerpts (all $ps < 0.05$).

### 6.2.3. Comparison with age-group mode

In order to investigate the effects of how consensus is defined on the results obtained, children's scores on all three tasks were recalculated as correct if they agreed with the modal choice of their age group as opposed to that of adults (Score B). Using this

scoring procedure, mean accuracy rose from 51.3% to 55.4% for 5-year-olds, from 62.3% to 70.2% for 7-year-olds, and from 74.9% to 78.5% for 9-year-olds. When these results were submitted to a 4(age) x 2(score) analysis of variance, significant main effects of age ( $F(3,115) = 61.96$, $p < 0.000$) and type of score ( $F(1, 115) = 79.45$, $p < 0.000$) were qualified by a significant Age x Score interaction ( $F(3, 115) = 13.42$, $p < 0.000$). Analysis of simple main effects for type of score indicated that the accuracy scores of 5-, 7- and 9-year-olds were all significantly greater when the children's responses were compared with their age-group mode as opposed to the adult mode (all $ps < 0.01$). Simple main effects for age were significant for both types of score. Follow-up analytic comparisons of means (Tukey's HSD test) with alpha set at 0.05 indicated that when responses were scored according to the adult mode (Score A), each age group was significantly more accurate than the age group directly below it (all $ps < 0.01$) When responses were scored according to the age group mode (Score B), the accuracy scores of 5- and 7-year-olds were still significantly lower than those of 9-year-olds and adults (all $ps < 0.001$). The revised accuracy scores of 9-year-olds, on the other hand, did not differ significantly from the accuracy scores of adults, suggesting that 9-year-olds were no less consistent than adults in their emotional categorisations of musical excerpts when their own age-group modal choice, rather than the adult mode, was taken as the standard or 'correct' judgment.

## 6.3. Base rates in the use of emotion labels

Research question 4 concerned the role played by children's and adults' basic inclinations to use, or not to use, certain emotion labels in their apparent accuracy for recognising these emotions. As Felleman et al. (1983) have pointed out, levels of overall accuracy when subjects are asked to give emotion labels to stimuli can be influenced by any systematic differences in the frequency with which different age groups use the different emotion category labels. Consequently, base rates for label usage were calculated by computing the mean number of times subjects used each of the emotion labels in Task 3 (the only task in which they were presented with all six options simultaneously) independent of whether they were used correctly or not.

These data, which are displayed in Table VI, were then subjected to 1-way ANOVAs for each emotion, with age as a between-subjects factor. Significant differences in base rates were found for the labels happy ( $F(3, 115) = 8.919$, $p < 0.000$), excited ( $F(3, 115) = 6.712$, $p < 0.001$) and scared ( $F(3, 115) = 5.670$, $p < 0.01$). Post-hoc planned comparisons (LSD test) indicated that adults labelled excerpts as happy significantly more often than all three groups of children, while 9-year-olds had higher base rates for using the label 'happy' than the 7-year-olds (all $ps < 0.01$). In addition, children in all three age groups labelled excerpts as scared more often than did adults (all $ps < 0.05$), while 7-year-olds used the label 'excited' significantly more often than all the other age groups (all $ps < 0.05$).

**Table VI.** Mean base rates (%) for the use of emotion labels in task 3 as a function of subject age

| Subject Age | Emotion label | | | | | |
|---|---|---|---|---|---|---|
| (years) | Calm | Happy | Excited | Scared | Angry | Sad |
| 5 | 19.1 | 15.5 | 14.1 | 19.4 | 14.1 | 17.8 |
| 7 | 22.2 | 12.5 | 22.2 | 16.3 | 12.1 | 14.7 |
| 9 | 21.9 | 17.7 | 17.5 | 16.5 | 10.2 | 16.2 |
| Adult | 21.3 | 21.7 | 17.6 | 11.8 | 12.6 | 15.0 |

### 6.3.1. Differential Accuracy

The above results indicating significant variation in the use of certain emotion labels suggest that children's accuracy scores for task 3 may have been reduced or enhanced because they tended to use certain labels more or less often than they were used by adults. To investigate this possibility, a differential accuracy score was calculated for each subject that controlled for the influence of base rates by dividing the number of times a subject used an emotion label correctly (defined as agreement with the adult norm, or Score A) by the total number of times they used that label (correctly or incorrectly). This score can be thought of as a conditional probability: given that a

subject used a particular label, what was the likelihood that the label was correct? Once computed, differential accuracy scores were compared to overall accuracy scores in a 4 (age) x 2 (type of score) ANOVA for each of the three emotion categories, with type of score treated as a within-subjects factor. If high base rates for using a particular label facilitated accurate recognition, overall accuracy should be significantly greater than differential accuracy, with the reverse being true if low base rates hindered recognition. Table VII presents a comparison of overall and differential accuracies (expressed as percentages) for the different age groups.

**Table VII.** Differential versus overall accuracy (% correct) for the use of emotion labels on task 3 as a function of subject age

| Subject Age (years) | Emotion label | | | | | |
|---|---|---|---|---|---|---|
| | Happy | | Scared | | Excited | |
| | Differential accuracy | Overall accuracy | Differential accuracy | Overall accuracy | Differential accuracy | Overall accuracy |
| 5 | 56 | 32 | 18 | 63 | 19 | 25 |
| 7 | 67 | 31 | 20 | 66 | 25 | 47 |
| 9 | 92 | 57 | 35 | 93 | 28 | 48 |
| Adult | 89 | 69 | 52 | 93 | 43 | 63 |

For happy excerpts, significant main effects of age ($F$ (3, 115) = 19.349, $p < 0.000$) and type of score ( $F$(1,115) = 144.235, $p < 0.000$) were qualified by a significant age × score interaction ( $F$(3, 115) = 2.811, $p < 0.05$). Planned analytic comparisons (LSD test) indicated that differential accuracy scores were significantly higher than overall accuracy scores for all age groups ( $p < 0.000$), suggesting that low base rates may have reduced subjects' accuracy for identifying happy excerpts. Post-hoc planned comparisons of the age-group means (LSD test) indicated that the overall accuracy scores of adults were significantly greater than those of 9-year-olds, who in turn performed with greater accuracy than the 5- and 7-year-olds (all $ps < 0.05$). The differential accuracy scores of the adults and 9-year-olds, however, did not differ

significantly from each other, although they were significantly higher than the differential accuracy scores of 7-year-olds, which in turn were higher than the differential accuracy scores of the 5-year-olds (all $ps < 0.05$). Thus when the children's relatively low base rates for using the label 'happy' (compared with adults) were controlled for, 9-year-olds and adults performed with approximately equal accuracy. Furthermore, 7-year-olds were more accurate than 5-year-olds, a difference that was not apparent in the overall accuracy scores.

Significant main effects of age ( $F(3, 115) = 10.065$, $p < 0.000$) and type of score ( $F(1, 115) = 285.091$, $p < 0.000$) were also found for the scared excerpt; there was no statistically significant interaction in this case. Overall accuracy scores were significantly higher than differential accuracy scores, suggesting that high base rates may have increased subjects' accuracy at identifying the scared excerpt. Post-hoc planned comparisons of means (LSD test, $\alpha = 0.05$) suggested that this was particularly the case for 9-year-olds: although their overall accuracy scores were equivalent to those of adults, their differential accuracy scores were significantly lower than those of adults ( $p < 0.01$), indicating that 9-year-olds were less accurate than adults when their relatively high base rates for using the label 'scared' were controlled for. The high base rates of 5- and 7-year-olds, on the other hand, did not appear to affect their accuracy: the performance of these two age groups was significantly poorer than that of 9-year-olds and adults for both overall and differential accuracy scores (all $ps < 0.05$).

For excited excerpts, significant main effects of age ( $F(3, 115) = 7.176$, $p < 0.000$) and type of score ( $F(1, 115) = 86.132$, $p < 0.000$) were again qualified by a significant interaction between age and type of score ( $F(3, 115) = 3.890$, $p < 0.05$). Post-hoc planned comparisons of means (LSD test, $\alpha = 0.05$) indicated that overall accuracy scores were significantly greater than differential accuracy scores for all age groups except the 5-year-olds, suggesting that high base rates for using the label 'excited' may have enhanced the older subjects' 'accuracy' for identifying those excerpts. The overall accuracy scores of adults were greater than those of 9- and 7-year-olds, who did not differ significantly from each other, but in turn were more

accurate than the 5-year-olds (all $ps < 0.000$). The differential accuracy scores of adults remained greater than those of all the other age groups, indicating that adults were still more accurate than the children for identifying excited excerpts when base rates were controlled for (all $ps < 0.01$). The only other finding that changed when differential accuracy scores were used is that there was no significant difference between the differential accuracy scores of 5- and 7-year-olds, suggesting that 7-year-olds were no more accurate than 5-year-olds when their relatively high base rates for using the label 'excited' were controlled for.

## 6.4. Intrasubject consistency

Hypothesis 3 stated that intra-individual consistency should increase with age. Subjects' responses to the music tasks were analysed for individual consistency or stability by giving each subject who completed all three tasks (n = 112) an inconsistency score indicating the mean number of categories into which they placed the eighteen excerpts across the three tasks (Score C). A one-way analysis of variance on this measure revealed that within-subject agreement increased significantly with age: mean inconsistency was 2.194 for 5-year-olds, 1.969 for 7-year-olds, 1.763 for 9-year-olds and 1.680 for adults[2] ( $F(3, 109) = 43.613$, $p < 0.000$). Post-hoc planned comparisons (LSD test) indicated that 5-year-olds were significantly more inconsistent than 7-year-olds, who were significantly more inconsistent than 9-year-olds and adults (all $ps < 0.0001$). Inconsistency scores of 9-year-olds and adults did not differ significantly.

## 6.5. Systematic inaccuracies

Analysing the data obtained in this study according to the traditional standard of agreement with the adult norm, it has been found that the inter- and intra-individual judgments of children (particularly those aged 7 or younger) regarding the emotional

[2] Inconsistency was not as great as these figures imply: these scores are inflated since the same emotion categories were not necessarily available to subjects on all three tasks. Nevertheless, even in cases where the same response options *were* available, even 9-year-olds and adults were not completely consistent in their judgments.

categorisation of musical excerpts are less consistent than those of adults. While many data analyses stop at this point, we also need to consider the possibility that children's (and adults') inaccurate judgments of the emotional character of the musical excerpts were not randomly distributed but rather represent systematic tendencies to 'confuse' one emotion category with another.

Initial indications that children's responses might be more orderly than the accuracy data suggest were obtained from chi square values calculated to compare actual frequency with which emotion labels were associated with musical excerpts against a frequency to be expected with random responding (6 out of a possible 18 for Task 1, 9 for Task 2, and 3 for Task 3). Overall, 216 goodness of fit chi square tests were carried out (18 musical excerpts, 4 age groups, 3 tasks) and 185 (86%) were significant with alpha set at 0.05. Adults' responses were different from a distribution expected by chance on 93% of the trials, 9- and 7-year olds' on 94% of the trials, and 5-year-olds' on 61% of the trials. Thus even where children did not agree with the adult standard (were 'inaccurate'), the majority of their responses were orderly.

To examine what the subjects were doing that was orderly, the distribution of responses for each age group in Task 3 was calculated across excerpts that expressed the same emotion according to the adult standard. These responses are expressed as percentages in Table VIII. As predicted by hypothesis 4, category membership was found to be graded: there were better and poorer musical exemplars of each emotion category. For calm excerpts, percentage agreement across age groups ranged from 37.9% (excerpt 12) to 67.2% (excerpt 6); for happy excerpts, from 38.6% (excerpt 2) to 56.2% (excerpt 5); for excited excerpts, from 26.7% (excerpt 17) to 64.9% (excerpt 1); for angry excerpts, from 27.9% (excerpt 15) to 58.45% (excerpt 7); and for sad excerpts from 57.8% (excerpt 11) to 76.42% (excerpt 3). Only one excerpt was classified by the majority of adults as expressing fear: mean agreement on this classification across age groups was 79.7%.

**Table VIII.** Distribution of responses (task 3) by age and emotion

| Adult modal emotion choice | Subject Age (years) | Distribution of emotion judgments | | | | | |
|---|---|---|---|---|---|---|---|
| | | Calm | Happy | Excited | Scared | Angry | Sad |
| Calm | 5 | **38** | 14 | 6 | 9 | 8 | 25 |
| | 7 | **54** | 7.3 | 4 | 5.3 | 0 | 29.3 |
| | 9 | **67** | 2 | 1 | 3 | 0 | 27 |
| | Adult | **68** | 9 | 1 | 1 | 0 | 21 |
| | Combined | **57** | 8 | 3 | 4.5 | 2 | 25.5 |
| Happy | 5 | 12 | **32** | 27 | 13 | 8 | 8 |
| | 7 | 8 | **33** | 48 | 4 | 6 | 1 |
| | 9 | 3 | **57** | 37 | 1 | 1 | 1 |
| | Adult | 0 | **67** | 31 | 0 | 2 | 0 |
| | Combined | 6 | **47** | 36 | 4.5 | 4 | 2.5 |
| Excited | 5 | 12 | 10 | **25** | 30 | 15 | 8 |
| | 7 | 5 | 8 | **48** | 27 | 12 | 0 |
| | 9 | 3.3 | 3.3 | **48** | 33.3 | 12 | 0 |
| | Adult | 3.3 | 3.3 | **63** | 17 | 13.3 | 0 |
| | Combined | 6 | 6 | **46** | 27 | 13 | 2 |
| Scared | 5 | 3 | 3 | 7 | **63** | 17 | 7 |
| | 7 | 0 | 0 | 6.67 | **66.67** | 26.67 | 0 |
| | 9 | 0 | 0 | 0 | **97** | 3 | 0 |
| | Adult | 0 | 0 | 7 | **93** | 0 | 0 |
| | Combined | .75 | .75 | 5 | **80** | 12 | 1.75 |
| Angry | 5 | 11 | 6 | 8 | 33 | **38** | 4 |
| | 7 | 3 | 2 | 10 | 38 | **47** | 0 |
| | 9 | 0 | 1 | 9 | 38 | **52** | 0 |
| | Adult | 0 | 1 | 9 | 27 | **63** | 0 |
| | Combined | 3.5 | 2.5 | 9 | 34 | **50** | 1 |
| Sad | 5 | 18.3 | 3.3 | 5 | 8 | 7 | **58.3** |
| | 7 | 37 | 3 | 0 | 3 | 0 | **57** |
| | 9 | 22 | 5 | 0 | 0 | 0 | **73** |
| | Adult | 18 | 0 | 0 | 0 | 0 | **82** |
| | Combined | 24 | 3 | 1 | 2.75 | 1.75 | **67.5** |

Percentages in bold type refer to choices that correspond to the adult modal choice

In an attempt to answer research question 5 and examine whether there were systematic age-related changes in modal emotion choices, the children's data were reviewed for evidence of instances in which the percentage of subjects within an age level who agreed on an alternative interpretation exceeded the percentage who agreed with the adult interpretation. In Task 3, this was found to be the case for 6 (although not exactly the same six) out of the 18 excerpts for 5-year-olds and 7-year-olds, and for 4 of the excerpts for 9-year-olds.

In order to examine whether these shifts in modal choices were systematic, age-group modal choices were subsequently compared with the adult standard across emotion categories, revealing that for excerpts identified by the majority of adults as happy, 48% of 7-year-olds' responses were excited interpretations, while only 33% were happy interpretations. For excerpts identified by the majority of adults as excited, 30% of the 5-year-olds' responses were scared interpretations, while only 25% were excited interpretations.

Hypothesis 5 predicted that the range of emotion categories selected would narrow with age. The data presented in Table VIII support this prediction. On average, adults' responses to the musical excerpts in Task 3 fell into 3 different categories (max. = 6). 9-year-olds' responses fell into a mean of 3.167 categories, 7-year-olds' into a mean of 4.11 categories, and 5-year-olds' into a mean of 5.56 categories. These values were found to differ significantly when examined by means of a one-way ANOVA ( $F(3, 68) = 30.72$, $p < 0.000$). Planned analytic comparisons (LSD test) indicated that 5-year-olds' responses fell into significantly more categories than 7-year-olds', whose responses in turn fell into significantly more categories than those of 9-year-olds and adults (all $ps < 0.01$). There was no significant difference in the number of categories utilised by 9-year-olds and adults.

Hypothesis 6 stated that for each musical excerpt there would be a focal point, the emotion category most likely to be chosen, with the probability of a choice decreasing with the distance of the emotion from the focal point according to the structural model. As we can see from the raw data presented in Appendix 4, it appears that for

the majority of musical excerpts there was a clear modal choice, the emotion label most likely to be chosen, but that in some cases this shifted with age. Moreover, if we look at Table VIII we can see that subjects' choices on task 3 were clustered in a systematic fashion. The six emotion categories in this table have been arranged in the same order as they appear in Russell et al.'s (1980, 1983, 1986) circumplex model of emotions. Thus, provided the two end emotions (sad and calm) are thought of as adjacent, the distribution of subjects' choices can be considered in terms of what would be predicted by the structural model. According to this model, the majority of errors (defined here as choices other than the adult modal choice) should fall into the two emotion categories that are adjacent to the 'correct' choice. By chance, only 40% of these errors would be the two adjacent emotions.

For the 5-year-olds, the mean percentage of errors across the 30 subjects that fell into adjacent categories was 60.5; the means for 7-year-olds, 9-year-olds and adults were 80.15%, 84.23% and 85.1% respectively. These means were analysed with one sample *t*-tests and found to be highly significantly different from chance ($p < 0.001$) for each age group.

Visual inspection of the data, however, suggested that this finding might in some cases simply reflect tendencies to confuse a particular emotion with only one other. 59.62% of errors on calm excerpts were sad interpretations, while only 18.46% were happy interpretations. 67.94% of errors on happy excerpts were excited interpretations, as opposed to only 10.79% calm interpretations. On excited excerpts, 49.61% of errors were scared interpretations, while only 11.63% were happy interpretations. On angry excerpts, 67.78% of errors were scared interpretations, while only 2.22% were sad interpretations. For sad excerpts, 73.08% of errors were calm interpretations, as opposed to only 5.13% angry interpretations. The only emotion category for which errors were more-or-less evenly distributed between the adjacent emotion categories was scared: 25% of errors were excited interpretations, and 20.83% were angry interpretations.

Furthermore, while the general pattern did show a modal response with declining probability of response to either side, there were a few minor exceptions to this pattern. 5- and 7-year-olds' 'misidentifications' of sad and calm excerpts were more likely to be scared judgments than the structural model would predict, while angry excerpts were more often misidentified by these two younger groups as calm than the model would predict.

## 6.6. Correlations with Hedonic Tone and Arousal ratings

Hypothesis 7 was largely supported by the finding already presented that with the exception of 5-year-olds' distinction of sad and calm excerpts, subjects of all ages were able to make reliable distinctions between musical excerpts varying on only one dimension, with the other underlying dimension held more-or-less constant.

Hypothesis 8 proposed that there would be a relationship between independent pleasure and arousal ratings of the music and the categories into which subjects placed them. In order to examine this relationship, mean pleasure ratings of the pilot subjects on Mehrabian and Russell's (1974) scales were correlated with the number of subjects in each age group who rated each excerpt as calm, happy or excited on task 3 (as opposed to sad, angry or scared), and mean arousal ratings were correlated with the number of subjects in each age group who labelled each excerpt as happy, excited, scared or angry (as opposed to calm or sad).

For pleasure ratings, correlations with the number of subjects classifying each excerpt as expressing positive emotions ranged from 0.69 for the 5-year-olds to 0.81 for the 9-year-olds; the correlation coefficient across all the age groups combined was 0.78. These were all significantly above chance levels (all $ps < 0.01$). $r$-sq ranged from 0.473 for the 5-year-olds to 0.663 for the 9-year-olds; taken across age groups, pleasure ratings accounted for 61.47% of the variance in the number of subjects associating the excerpts with positively valenced emotions on task 3.

Correlations between arousal ratings and the number of subjects categorising each excerpt as expressing active emotions ranged from 0.89 for the 5-year-olds to 0.94 for the adults and 9-year-olds, and all were highly significant (all $ps < 0.000$). $r$-sq values ranged from 0.80 for the 5-year-olds to 0.885 for the adults; combined across age groups, arousal ratings accounted for 88.4% of the variance in the number of subjects associating each excerpt with active as opposed to passive emotions on task 3.

## 6.7. Basis for emotion judgments of musical excerpts

Research question 6 was concerned with the basis for subjects' emotional categorisations of the musical excerpts. Subjects' spontaneous comments and answers to direct questions suggested that listeners of all ages used two main 'techniques' for assigning musical excerpts to emotion categories.

The most common technique across age groups was a form of 'visualisation' - forming 'mental pictures' of movie or stage scenes ("like a pirate movie" [5-yr-old]; "kissing movie music" [7-year-old]; "where the dinosaurs die" [7-year-old]) or of cartoons, dance motions, or other institutionalised occasions or settings such as circuses and formal or religious ceremonies ("Academy Award music" [Adult]; "sounds royal...they might be getting married" [5-yr-old]; "sounds like Christmas" [5-year-old]). Some listeners also pictured more personal or idiosyncratic scenes (e.g. "like a big monster stepping" [5-year-old]; "a herd of cheetahs and lions" [5-year-old]).

All age groups also used the formal musical elements of the excerpts as a basis for their judgments, although the adults tended to describe these in more technical terms. Even the 5-year-olds made reference to tempo: for example, " 'cause it's going fast"; and "very quick" were justifications given for describing excerpt 1 as excited. Subjects of all ages also made reference to the rhythm of the musical excerpts. The latter appeared to be particularly important for excerpts judged as angry: while adults

referred to "deep drum sounds", 5-year-olds couched their responses in terms such as "boom boom" (excerpt 14). Adults also made reference to other musical elements such as articulation (staccato versus legato), melodiousness, pitch level, instrumentation and modality, the last 3 of which were claimed to be useful for making distinctions between excerpts similar in activity level. In addition, both adults and children made more general and ambiguous comments about features of the music such as "Sad music has a 'crying' sound" (adult) and "If it goes woo, woo, woo, then I know it's sad" (5-year-old). A few adults commented that formal musical elements were particularly useful for making subtle distinctions between closely related emotions such as happiness and excitement or fear and anger. An analysis of the quantitative data obtained was subsequently used to examine the effects of these musical elements in more depth.

The quantitative part of the statistical analysis examined the relationship between subjects' responses on task 3 and the formal musical characteristics of the excerpts on the stimulus tape. The means of the judges' ratings for tempo, rhythm (firm vs. flowing), rhythmic activity, articulation, pitch range and harmony as well as the codings for modality were initially correlated with the number of subjects in each age group (as well as with the mean across age groups) who placed each musical excerpt in each of the six emotion categories. That is, the mean ratings of the musical elements were correlated with the number of calm judgments, the number of happy judgments, the number of sad judgments, and so on. They were then correlated with the number of subjects in each age group, and with the mean across age groups, who responded to each excerpt with calm, happy or excited judgments (pleasant emotions) as opposed to the unpleasant emotions of sadness, fear and anger, as well as with the number who responded with 'active' emotion labels (happy, excited, scared or angry) as opposed to the passive emotions of sadness and calmness. The results of these correlation matrices for the mean number of subjects across age groups are presented in Table IX; full results for each age group separately are displayed in Appendix 6.

**Table IX.** Correlations between musical elements and emotional categorisations of mean number of subjects across age groups combined (task 3)

| Musical elements | Emotion | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | CALM | HAPPY | EXCITED | SCARED | ANGRY | SAD | PLEASANT | ACTIVE |
| TEMPO | -0.70* | 0.41 | 0.83* | 0.35 | 0.12 | -0.75* | 0.23 | 0.85* |
| RHYTHM | -0.82* | 0.34 | 0.65* | 0.37 | 0.54* | -0.78* | -0.01 | 0.94* |
| RHYTHMIC ACTIVITY | -0.61* | 0.30 | 0.74* | 0.40 | 0.20 | -0.79* | 0.18 | 0.82* |
| ARTICU-LATION | -0.56* | 0.38 | 0.53* | 0.43 | 0.31 | -0.84* | 0.13 | 0.82* |
| HARMONY | -0.38 | -0.62* | -0.24 | 0.79* | 0.59* | -0.07 | -0.82* | 0.27 |
| MODALITY | 0.03 | 0.57* | 0.52 | -0.41 | -0.30 | -0.70* | 0.92* | 0.41 |

* Indicates significant at $\alpha = 0.05$

As can be seen from the data presented in Appendix 6, the patterns of correlations were very similar across age groups. Consequently, stepwise multiple regression analyses conducted to explore these relationships in more detail were calculated using the mean number of judgments across age groups for each emotion category as the criterion variable, and the judges' ratings of the six musical elements as predictor variables.

**Table X.** Intercorrelations of musical elements

| | TEMPO | RHYTHM | RHYTHMIC ACTIVITY | ARTICULATION | HARMONY |
|---|---|---|---|---|---|
| TEMPO | | | | | |
| RHYTHM | 0.81* | | | | |
| RHYTHMIC ACTIVITY | 0.95* | 0.81* | | | |
| ARTICULATION | 0.77* | 0.75* | 0.82* | | |
| HARMONY | 0.12 | 0.21 | 0.15 | 0.17 | |
| MODALITY | 0.53 | 0.27 | 0.56* | 0.60* | -0.67* |

* Indicates significant at $\alpha = 0.05$

The number of subjects describing excerpts as calm was highly correlated with rhythm for all age groups (all $rs \leq -0.75$), and was also moderately and significantly ( $p < 0.05$) correlated with tempo, rhythmic activity and articulation: correlation coefficients for these elements ranged from -0.49 to -0.76. As can be seen in Table X, these variables were also highly correlated with each other.

A stepwise multiple regression analysis was subsequently performed with the mean number of calm judgments across age groups as the criterion variable and the six musical elements as the predictor variables. The results of the regression model obtained indicated that the predictor variables of rhythm, harmony, rhythmic activity and tempo, when taken together, accounted for 76% of the variance in the number of subjects describing the music as calm ( $F(4, 13) = 14.478$, $p < 0.00$). Rhythm was the single best predictor, accounting for approximately 67% of the variance in subjects' responses ( $t(13) = -3.57$, $p < 0.01$, $\beta = -0.75$.) These results suggested that excerpts with flowing rhythms, consonant harmonies, little rhythmic activity and slow tempos tended to be judged as calm. The absence of articulation from this regression model is notable, considering its high simple correlation with the criterion variable. However, articulation was quite highly correlated with rhythmic activity, rhythm and tempo, and the relationship between calm judgments and legato articulation apparent from the simple correlations was masked in the regression analysis because these elements were more strongly related to the criterion variable.

The second correlation matrix showed that the number of subjects in each age group describing excerpts as happy was significantly correlated with harmony ($-0.63 \leq$ all $rs \leq -0.57$), and also showed a moderate correlation with modality ($0.47 \leq$ all $rs \leq 0.65$), although this relationship only reached statistical significance ( $p < 0.05$) for the 7-year-olds and adults. Harmony and modality were also significantly correlated with each other ($r = -0.67$).

A stepwise regression analysis indicated that 68% of the variance in the mean number of subjects across age groups responding with happy judgments could be predicted from a combination of harmony, articulation, tempo and rhythmic activity ( $F(4, 13)$

= 10.17, $p < 0.00$). Excerpts tended to be judged as happy if they were characterised by consonant harmonies, staccato articulation, fast tempos and much rhythmic activity. The absence of modality from the regression equation despite its high simple correlation with the number of happy judgments can be attributed to its high correlation with harmony: when harmony was omitted as a predictor variable, modality entered the regression model at the first step and accounted for approximately 26% of the variance in the number of happy responses. This suggests that excerpts rated as happy also tended to be in the major mode.

The number of subjects identifying excerpts as excited showed a significant correlation with tempo, rhythmic activity and rhythm for all age groups ($0.56 \leq$ all $rs \leq 0.85$), as well as a moderate correlation with articulation ($0.45 \leq$ all $rs \leq 0.55$) which reached statistical significance ($\alpha = 0.05$) for the 5-year-olds and adults. Multiple regression results indicated that tempo alone accounted for nearly 58% of the variance in the number of subjects responding with the label 'excited'; harmony was the only other predictor variable which contributed significantly towards *R*-sq, adding another 0.10 to its value ( $F(2, 15) = 18.86$, $p < 0.00$). Thus, fast excerpts with consonant harmonies tended to be judged as excited. The relationship between excited judgments and large amounts of rhythmic activity and staccato articulation which are apparent in the simple correlations did not appear in the regression analysis as they were masked by these variables' high correlations with tempo.

The only musical element with which the number of scared judgments was significantly correlated was harmony, with correlation coefficients ranging from 0.68 to 0.85 for the different age groups. A stepwise regression analysis indicated that this was indeed the most significant predictor of the number of scared judgments ( $t(15) = 5.40$, $p < 0.00$, $\beta = 0.74$); however, articulation also contributed significantly towards *R*-sq ( $t(15) = 2.25$, $p < 0.05$, $\beta = 0.31$). These two elements combined accounted for 69% of the variance in the numbers of scared judgments ( $F(2, 15) = 19.7$, $p < 0.00$), suggesting that a label of 'scared' tended to be applied to dissonant and staccato music.

Harmony was also significantly correlated with angry judgments for all age groups ($0.51 \leq$ all $rs \leq 0.69$), while moderate correlations between the number of angry responses and rhythm were significant for all ages with the exception of the 9-year-olds ($0.43 \leq$ all $rs \leq 0.59$). These two variables entered the stepwise regression analysis first, together accounting for approximately 72% of the variance in the number of angry responses. Inclusion of tempo and modality into the equation increased $R$-sq significantly, with all four variables together accounting for 78% of the variance ( $F(4, 13) = 15.85, p < 0.00$). These results suggest that subjects tended to describe excerpts with dissonant harmonies, firm rhythms, fast tempos and in the minor mode as sounding angry.

The number of sad categorisations in each age group was significantly correlated with all of the musical elements other than harmony, with correlation coefficients ranging between 0.59 and 0.87. Results of a stepwise multiple regression analysis indicated that articulation, modality and rhythm together accounted for 83% of the variance in the number of sad judgments ( $F(3, 14) = 28.67, p < 0.00$): in other words, legato excerpts with a flowing rhythm and in the minor mode tended to be described as sad. The association between sad judgments and slow tempos as well as little rhythmic activity suggested by the simple correlations was not apparent in the regression model as these elements were highly correlated with rhythm and articulation, which were more strongly related to the number of sad judgments.

The number of subjects describing the excerpts as expressing pleasant emotions was found to be highly and significantly correlated with harmony and modality for all age groups, with correlation coefficients ranging between $\pm 0.72$ and $\pm 0.95$. Results of a stepwise multiple regression analysis indicated that harmony was the single most significant predictor ( $t(13) = -5.79, p < 0.000, \beta = -6.95$), accounting for approximately 68% of the variance on its own. Together with modality, tempo and rhythm it accounted for approximately 90% of the variance in responses ( $F(4, 13) = 17.45, p < 0.00$), suggesting that subjects tended to associate excerpts with pleasant emotions particularly if they had consonant harmonies and were in the major mode,

while faster tempos and flowing rhythms were also associated with more pleasant emotion states.

The number of subjects categorising excerpts as expressing active emotions was significantly correlated with rhythm, tempo, rhythmic activity and articulation ($0.74 \leq$ all $rs \leq 0.94$), and again this pattern of relationships was apparent for all age groups. A stepwise multiple regression analysis indicated that rhythm was the most important single predictor variable ( $t\,(12) = 5.45$, $p < 0.00$, $\beta = 0.64$). Together with articulation, tempo, rhythmic activity and harmony it accounted for approximately 93% of the variance in subjects' responses ($F\,(5, 12) = 46.927, p < 0.000$). A firm (as opposed to flowing) rhythm therefore appeared to be the most important musical element for predicting the number of subjects associating the excerpts with active emotions; however, active judgments were also related to staccato articulation, fast tempos, much rhythmic activity and dissonant harmonies.

## 6.8. Effects of personal characteristics

Although there was a general tendency for consistency of emotional judgments to increase with age, there were marked individual differences in accuracy (for our purposes, agreement with the adult mode, or Score A). The highest scoring 5-year-old, for instance, was substantially more accurate than the lowest scoring adult (70% versus 57% correct).

As an initial step in investigating what variables might be related to individual differences within age groups and attempting to answer research question 7, t-tests were used to investigate whether there were differences in the mean accuracy across the three music tasks of males and females within each age group. Accuracy was defined as the number of responses that agreed with the adult modal choice out of a total of 54 (18 musical excerpts each judged 3 times). Only one significant difference was found: 9-year-old females (M = 42.33) were significantly more accurate ( $t\,(27) = -2.075, p < 0.05$) than 9-year-old males (M = 38.765).

In order to check these results, subject sex was included in correlation matrices for each age group along with the following independent variables: control task score (all age groups), reading and maths scores (7- and 9-year-olds only) and number of years formal musical training (adults only). Once again, the dependent variable was the accuracy score (compared with the adult standard) across all three music tasks combined.

As far as sex is concerned, the correlation matrices tended to confirm the results of the t-tests. For 5-year-olds, there was a fairly low and nonsignificant ($\alpha = 0.05$) correlation ($r = -0.33$) between sex and accuracy on the music tasks in favour of boys, which accounted for 11.06% of the variance in subjects' performance. There was also a slight tendency for boys to perform better than girls on the control task, although here the relationship was somewhat smaller ($r = -0.23$, $r$-sq = 0.053). For 7-year-olds, there was a very small and nonsignificant correlation between sex and accuracy on the music tasks, again in favour of boys ($r = -0.1$, $r$-sq = 0.011), and a slightly higher but still nonsignificant correlation between sex and the control task score in favour of girls ($r = 0.33$, $r$-sq = 0.11). The accuracy of adults on the music tasks showed only a very slight and nonsignificant correlation with sex in favour of females ($r = 0.12$, $r$-sq = 0.015), and there was virtually no relationship between sex and the control task score ($r = 0.07$). As expected from the t-test results, a stronger, although still fairly small and statistically nonsignificant relationship between sex and accuracy on the music tasks in favour of girls was found for 9-year-olds ($r = 0.33$, $r$-sq = 0.11). When sex was correlated with accuracy scores for each emotion separately, however, it was found that the 9-year-old females' advantage was limited to the identification of happy excerpts ($r = 0.58$, $r$-sq = 0.34, $p < 0.05$). Sex showed only very low and nonsignificant correlations with accuracy for the other emotion categories. Control task scores also showed a weak and nonsignificant correlation with sex in favour of girls ($r = 0.19$, $r$-sq = 0.036).

A closer look at the results indicated that for 9-year-olds (but not 7-year-olds) sex was also moderately, and significantly, correlated with reading and maths scores (reading: $r = 0.42$, $r$-sq = 0.18, $p < 0.05$; maths: $r = 0.48$, $r$-sq = 0.23, $p < 0.01$), which were

themselves moderately intercorrelated ($r = 0.45$, $r$-sq = 0.20, $p < 0.05$) and in turn showed a positive, albeit rather low and nonsignificant correlation with accuracy on the music tasks (reading: $r = 0.25$, $r$-sq = 0.062; maths: $r = 0.11$, $r$-sq = 0.01). The 7-year-olds' reading and maths scores were highly intercorrelated ($r = 0.82$, $r$-sq = 0.68, $p < 0.001$); however, only maths scores showed a (nonsignificant) correlation with sex in favour of girls ($r = 0.33$, $r$-sq = 0.11). Reading scores, but not maths scores, were slightly positively correlated with accuracy on the music tasks ($r = 0.20$, $r$-sq = 0.04), although again these correlations did not reach statistical significance.

Subjects' scores on the control task showed only low and nonsignificant ($p > 0.05$) correlations with accuracy on the music tasks for all age groups. The highest correlation coefficient, $r = 0.26$, was found for the preschoolers; for older children and adults the correlations ranged from 0.14 to 0.16. Thus, subjects' accuracy on the control task accounted for only between 2 and 7% of the variance in their accuracy on the music tasks.

As only 5 of the 90 children were known to have had any formal musical training, correlations between the amount of musical training and accuracy on the music tasks were only computed for adults. The number of years of musical training, which was unrelated to sex, showed only a low and nonsignificant positive correlation with accuracy on the music tasks ($r = 0.19$, $r$-sq = 0.04) and, curiously, a slightly higher , but still nonsignificant, correlation with performance on the control task ($r = 0.25$, $r$-sq = 0.06).

# CHAPTER 7: DISCUSSION

This study has identified several broad trends in the development of children's interpretation of emotion in music which are consistent with the findings of previous research. The ability to identify emotions in music was found to emerge in the preschool years and to improve with age. In addition, the consistency of subjects' judgments was found to vary according to the particular emotion being expressed, and subjects' 'misjudgments' (that is, their selection of emotion labels other than the adult modal choice) were found to be systematic. However, the results also provided some contributions that are unique, and others that conflict with the existing literature. Furthermore, even though the majority of the findings are in accordance with those of previous work, the interpretations provided for these findings will differ somewhat from the interpretations that have been provided in the past.

## 7.1. Effects of age

The finding that even preschool children generally show beyond-chance agreement with each other and with adults in judging the emotional character of short, decontextualised musical excerpts supports the conclusions of previous studies which indicate that children's sensitivity to emotional expression in an abstract stimulus like music follows closely in time on their ability to identify more conventional, everyday expressions of emotion found in the human face and voice (Cunningham & Sterling, 1988; Dolgin & Adelson, 1990; Meerum Terwogt & van Grinsven, 1988, 1991; Trehub et al., 1987 in Trehub, 1993). Despite the fact that the ability to identify emotions in music does not have the same ecological advantages or social consequences as the ability to identify emotions expressed by one's fellow humans, the general finding of these studies is that the ability to apply emotion terms to music in a reliable manner is well established by 6 or 7 years of age. This conclusion was supported by the findings of this study, even though the range of possible emotion

choices was increased from the 2 - 4 used in previous studies to 6, and despite the fact that the musical excerpts were not preselected as prototypical examples of particular emotion categories. This suggests that (as other recent research on children's understanding of emotion has indicated) young children have a more advanced understanding of emotional expression than they have traditionally been credited with. Furthermore, the ability to identify emotion in music appears to be little influenced by a person's gender or intellectual skills - a conclusion that has also been reached by several other researchers (Behrens & Green, 1993; Campbell, 1942; Hevner, 1935; Meerum Terwogt & van Grinsven, 1991), although Beldoch (1964) did find that college students' verbal (both not mathematical) aptitude was related to their ability to interpret emotions in music.

This study is also consistent with the previous literature on children's interpretation of emotion in music - as well as in more naturalistic stimuli such as facial expressions - in finding that children's agreement with the modal emotion choice of adults tends to increase with age, particularly up to the age of nine or thereabouts (Cunningham & Sterling, 1988; Dolgin & Adelson, 1990; Meerum Terwogt & van Grinsven, 1988, 1991; Scherer & Oshinsky, 1977). What these studies typically fail to do, however, is to provide any systematic explanation of how and why these age-related changes occur.

One possibility is that children simply make more errors than adults do due to lack of concentration, confusion, a shorter attention span, a tendency to be distracted, and so on. Such an explanation cannot be completely dismissed: on the whole the adult sample in this study did appear to concentrate more closely on the tasks, to listen more carefully to the excerpts and to put a large amount of thought into their judgments, whereas children were more likely to respond instantaneously. Nevertheless, the results of this investigation suggest that such an explanation is far from sufficient to account for the greater agreement with the adult standard in the emotion judgments made by older subjects.

## 7.2. Effects of musical elements

Another possible explanation that has been raised for these findings is that age-related increases in consistency may be related to an improved ability to perceive or differentiate the relevant acoustic emotion cues. The results of this study do not, however, support this explanation. Firstly, although the influence of tempo, rhythm and articulation on emotion judgments were similar to the role these elements have been found to play in expressing emotion in speech (Scherer, 1986; Sundberg, 1982), accuracy on the control task showed little relationship to accuracy on the music tasks. Although the control task was only a brief measure whose validity for measuring subjects' ability to identify emotions expressed in speech has not been established, these results do suggest that sensitivity to emotion in music does not depend solely, or even largely, on the ability to perceive auditory 'emotion detectors' common to both speech and music. Rather, it appears to require an awareness of features that are in the music itself, and not simply or exclusively extra-musical features of the voice.

Secondly, the patterns of correlations between musical elements and emotional categorisations found in this study were remarkably similar for all age groups, which is inconsistent with Sloboda's (1985) hypothesis that while all subjects should be able to make broad emotional distinctions between musical selections based on global, iconic characteristics like tempo and rhythm, only those who possess a fair amount of musical sophistication will be able to perceive and make discriminations based on more subtle (symbolic) aspects of the music such as tonal relations. Although the multiple regression results need to be interpreted with caution due to the high intercorrelations among musical elements, when viewed in conjunction with the simple correlations they contradict traditional beliefs that children cannot perceive the difference between consonant and dissonant harmonies until they are around 8-10 years old (Moog, 1976a in Kastner & Crowder, 1990). They are also inconsistent with Trehub et al.'s (1987) finding (cited in Trehub, 1993) that only musically trained adults made the conventional emotional associations to short melodic sequences in the major and minor modes. Rather, they support the findings of Crowder (1995; Kastner

& Crowder, 1994) and Kratus (1993) in suggesting (although not proving) that - at least in the context of other musical elements - even young children are able to perceive quite subtle aspects of music such as harmony and modality, and that they associate these and other musical elements with emotions in the same way that adults do. Although we cannot rule out the possibility that subjects' judgments were influenced by musical elements that were not investigated in this study, there was no evidence that children were focusing on different or fewer elements than those used by adults, nor that they were giving different weightings to the same features.

In fact, although adults were able to describe the musical elements in more technical terms than children, subjects of all ages appeared to make use of a fairly standard set of musical features which they interpreted in a way that was largely consistent with the findings of previous research (Crowder, 1985, 1994; Hevner, 1935, 1936, 1937; Kratus, 1993; Nielzén & Cesarec, 1982; Rigg, 1940; Scherer & Oshinsky, 1977; Wedin, 1972). The only finding that was not in accordance with the results of previous research was the slight tendency for children (but not adults) to rate excerpts with consonant harmonies as excited. This finding may reflect the relatively ambiguous position of excitement on the dimension of hedonic tone for adults: perhaps dissonant harmonies are associated with a more unpleasant, tense sort of excitement and consonant harmonies (which were correlated with judgments of pleasant emotions by all age groups) with a more positive, joyful form of excitement.

The quantitative results of this study also provide some support for the hypothesis put forward by Francés and Bruchon-Schweitzer (1983) that musical variables related to activation are iconic (i.e. similar to those that reflect activation in the voice and in body movement), while evaluative distinctions rely more on conventional, intrinsically 'musical' features like harmony and tonality. Reasons given by the subjects for their judgments generally support this conclusion, but suggest that it is not a hard-and-fast rule: some music was judged angry because the rhythm resembled a person stamping; some was judged sad because it sounded like a person crying. The distinction between iconic and more conventional features of music is not, however, always clear-cut, and at times the two are difficult to separate. In addition, it needs to

be remembered that other potentially important musical elements such as volume, timbre and various melodic characteristics were not investigated in this study; nor were the interaction effects of musical elements. These remain a potential topic for future research.

Regardless of whether musical elements connote emotion iconically or in a more conventional, symbolic fashion, however, the results of this study suggest that an advanced or sophisticated knowledge of music is not necessary for one to be able to interpret it in emotional terms. Additional support for this conclusion is provided by the finding that formal musical training was unrelated to adults' 'accuracy' at identifying the emotions expressed in music.

This does not, of course, discount the possibility that adults may be more consistent in their emotional interpretations than children because they have more experience with musical conventions or 'metaphors' - in other words, with the indexical associations of music. Sloboda (1985) refers to this as musical acculturation (or enculturation): knowledge about music which is acquired spontaneously, without the self-conscious, directed efforts to improve musical skills that constitute musical training. Although Meerum Terwogt and van Grinsven (1988) found that experience with music - both active (performance) and passive (the importance one attaches to music and the frequency with which one listens to it) had no significant effect on their subjects' ability to identify the emotions expressed in music, they point out that a small sample size and limited within-group differences in experience might have obscured any effects of the latter. Indeed, reasons given by the subjects in this study for their answers do suggest that the ability to interpret emotions in music is influenced by a knowledge of the types of music that typically accompany particular kinds of scenes or emotional episodes in movies, ballets, cartoons and the like, as well as of the types of music that are regularly associated with particular occasions, settings, or ceremonies. Some adults even spoke about a "film sense" and a "conditioned response to movie music". To what extent this knowledge depends on age, however, is a matter for conjecture: even the youngest subjects in this study seemed to have a good idea about what kinds of music tend to accompany love scenes, death scenes,

chases and so on in the media. Some of the 5-year-olds made specific references to movies from which they believed the excerpts came, and although these were inaccurate except for Pinocchio, some of their responses were remarkably perceptive. Examples are the 5-year-olds who described excerpt 5 ('Hero's Wedding') as sounding "royal...they might be getting married", and excerpt 14 ('The Fall of Saigon') as depicting "mad soldiers - cross soldiers".

## 7.3. Effects of emotion

The results of this study therefore suggest that increasing knowledge of musical conventions does not go a very long way towards explaining age-related differences in the consistency with which individuals interpret emotion in music. Rather, they raise the possibility that an improved ability to understand emotional expression in music may reflect not so much an increase in musical sophistication as an increase in knowledge about and understanding of emotion. As Barwell (1986) has stated, in order to understand what it means to call music sad or to be able to recognise the sadness in the music, one first has to know what it means to say that a person is sad, and to be able to recognise his or her sadness.

That this is in fact the case is suggested by previous studies which indicate that children find some emotions easier to identify in music than others (Cunningham & Sterling, 1988; Dolgin & Adelson, 1990; Meerum Terwogt & van Grinsven, 1991). More specifically, children apparently learn to identify happy music, facial expressions and situations first, with knowledge of more complex or unfamiliar emotions gradually being acquired later.

Viewed from this perspective, the results of the present study are puzzling. In contradiction to previous research, given a choice of all six emotion categories (task 3), children - like adults - identified the scared excerpt 'correctly' most often, followed by the sad excerpts. Performance on happy excerpts was relatively poor in comparison.. On Task 1 (Activation dimension varied) sad and calm excerpts were identified correctly most often, while on Task 2 (Pleasure Dimension varied) these

two emotion categories were identified correctly *least* often. How do we explain these discrepancies?

### 7.3.1. Musical stimuli

One possibility is that the results obtained for Task 3 reflect the particular musical excerpts used: perhaps the scared and sad excerpts were particularly good examples of music expressing fear and sadness, while the happy excerpts were only mediocre examples of this emotion category. This might have been the case particularly as (unlike in most previous research with the exception of Kratus, 1993) - the scoring procedures used in this study meant that the excerpts were not equally divided among emotion categories. In fact, only one excerpt was labelled as scared by the majority of adults and two as sad, whereas five were classified as happy.

This unequal distribution of musical excerpts among emotion categories does mean that the findings regarding the effects of the emotion being expressed on subjects' 'accuracy' need to be interpreted with caution. It is certainly true that excerpt number 10 ('Whale Chase') may have been a particularly good example of fear in music: four out of the five pilot subjects included synonyms of 'afraid' in their open-ended emotional descriptions of the music, whereas in previous studies (Cunningham & Sterling, 1988; Meerum Terwogt & van Grinsven, 1991) the percentage of pilot subjects agreeing on the emotional categorisation of scared excerpts was somewhat lower than the percentage agreeing on the emotional character of the other excerpts. On the other hand, the pilot subjects' descriptions do not suggest that the 'sad' excerpts used in this study were particularly good examples of this category, nor that the 'happy' excerpts failed to express this emotion clearly. Furthermore, if the consistency of subjects' responses was determined by how clearly particular emotions were expressed in the musical excerpts, we would expect that particular emotion categories would be identified correctly more often on all three tasks, which was not the case. Therefore while the particular selection of musical excerpts may have contributed to the results of this study, this is not a sufficient explanation for the present findings.

### 7.3.2. Base rates

Another possibility that was examined in this study is that agreement among subjects with the adult modal choice may be influenced by factors unrelated to their ability to interpret emotional expressions in music, and hence be inadvertent. The effects of base rates for using certain emotion labels on accuracy was felt to be a particularly important consideration in this study because (as discussed above) the musical excerpts were not evenly divided among the six emotion categories. The present results help to account for the contrasting findings of this study and previous research regarding the effects of emotion by suggesting that 'accuracy' for the identification of scared and excited excerpts in task 3 may have been enhanced by subjects' relatively high base rates for using those labels, while accuracy for the identification of happy excerpts may have been reduced by relatively low base rates for using the happy label.

The results of this study also indicate that in a few instances, 'accuracy' and 'inaccuracy' in identifying the emotional character of musical excerpts in Task 3 was in fact significantly influenced by adults' and children's varying preferences for the use of particular emotion terms. Happy excerpts, in particular, were not as difficult for the 7- and 9-year-old children in this study to interpret as their low overall accuracy rates, compared to those of adults and 5-year-olds respectively, might suggest: 9-year-olds performed equivalently to adults, and 7-year-olds performed with greater accuracy than 5-year-olds, when their low base rates for using the label 'happy' (compared with adults) were controlled for. The low base rates of 7-year-olds in relation to 9-year-olds did not, however, appear to affect the relative accuracy of these two groups: 9-year-olds were more accurate at identifying happy excerpts than 7-year-olds even when their differential base rates were controlled for. In addition, fear in music was not as easy for the 9-year-old children to recognise as their overall accuracy rates for Task 3 appear to indicate: when the children's relatively high base rates for using the 'scared' label were controlled for, 9-year-olds' accuracy fell significantly below that of adults. Similarly, the 7-year-olds' high base rates for using the 'excited' label apparently increased their accuracy at recognising that state

relative to the 5-year-olds: when base rates were controlled for, they did not perform significantly better than the preschoolers.

Only one previous study (Cunningham & Sterling, 1988) has taken children's base rates for using certain emotion labels into account. Although these authors also found that their youngest subjects (4-year-olds) had high base rates for scared responses, this did not appear to influence their accuracy. Nevertheless, the results of the present investigation support the point made by Felleman et al. (1983) that research on children's understanding of emotional expressions must recognise that 'accuracy' or 'inaccuracy' in identifying expressions may be influenced by more than the mere ability to interpret expressive cues 'correctly'.

*Why* children and adults use different emotion labels with different frequencies is a question that still needs to be investigated. Is fear simply more common in children than in adults? Possibly, but it seems doubtful that happiness is less common. Perhaps, however, these findings partly reflect the subjects' ways of responding to the task. It was observed during the experimental trials that when unsure of how to respond, the children who participated in this study tended to look for patterns in their answers (such as choosing the first emotion in block 1, the second in block 2, the third in block 3, and so on), or to choose an emotion that they felt had been too long neglected. Adults, on the other hand, tended to be less concerned about the frequencies with which they used the different emotion labels, and simply to select the term which they felt was most appropriate to describe a particular musical excerpt.

While the particular selection of musical excerpts and subjects' base rates for choosing particular labels go some way towards explaining why some emotions were identified more consistently than others, the evidence obtained in this study also suggests an alternative interpretation. It appears that (a) the varying effects of emotion for the different tasks, as well as (b) the discrepancies between this study and previous findings concerning the differential accuracy for different emotion categories and systematic 'confusion' of particular emotions and (c) the broad age-related trends observed, can all be accounted for if we shift our focus away from simply examining

children's accuracy in comparison to the adult standard (assuming that this represents the 'correct' way to describe each excerpt), and instead focus on the meaning and nature of emotion concepts at different ages from the perspective of prototype theory.

## 7.4. Prototype analysis

The findings of this study are consistent with the idea that emotion concepts - as used in everyday life - are prototypically organised and have fuzzy boundaries. For all age groups, some musical excerpts belonged quite clearly in a particular emotion category. Others, however, were more ambiguous and difficult to categorise, and were inconsistently described as expressing two or more different emotions: no sharp boundary separated excited musical excerpts, for example, from scared excerpts. Thus for each emotion category, some excerpts were good exemplars, others were poorer, and one excerpt could be a clear member of one category (e.g. happy) and a peripheral member of another (e.g. excited). These findings support the data obtained by Bullock and Russell (1984, 1986; Russell & Bullock, 1985, 1986) on children's and adults' categorisation of emotional facial expressions.

The present results also indicate that although category membership was graded and category boundaries fuzzy for all age groups, children's categories tend to become less fuzzy as they grow older. This was shown in three different ways.

Firstly, consensus among subjects about the emotion expressed in each excerpt tended to increase with age (at least up to the age of nine) regardless of whether consensus was defined as agreement with the adult or the age-group modal emotion choice. This finding supports Nelson and Nelson's (1978) hypothesis (cited in Alexander & Enns, 1988) that decisions about category membership become more conventional (that is, increase in social consensus) as people grow older.

Secondly, there was an age-related increase in the stability of category boundaries within individuals, again at least up until the age of 9. Because children were tested on two or three different days and adults at a single sitting, it might be argued that this

finding merely reflects procedural differences: children may simply have had more difficulty than adults in remembering the excerpts and their previous responses to them because of the greater time span separating tasks and their more limited memory capacity. However, the fact that the 9-year-olds were more consistent than the 7-year-olds despite the same procedure being followed for the two groups, and that 9-year-olds were no less consistent than adults despite experiencing a greater delay between tasks, tends to argue against this explanation.

Therefore while a degree of caution is advisable in interpreting these findings, they do suggest that for 5- and 7-year-olds the fuzzy category boundaries seen in the group data were a result of their individual consistency across tasks as well as their lack of agreement with one another within a task, whereas the fuzzy category boundaries seen in the 9-year-old and adult group data were primarily the result of disagreement among individuals (although some intra-individual inconsistency remained). A similar age-related trend was found in Alexander and Enns' (1988) study of children's categorisation of a set of toy animals whose group membership was ambiguous, although in their study individuals achieved intra-individual consistency at an earlier age.

This question of intra-individual consistency is one which has not previously been investigated with regard to children's interpretation of emotion in music, despite the possibility of doing so: in the studies by Cunningham and Sterling (1988) and Kratus (1993) subjects responded to the same excerpts twice. The findings of this investigation suggest that these researchers' failure to examine the consistency of individual children's judgments is an important oversight, as examining the stability of individual judgments gives us information about changes in children's emotion concepts which is not obtained by simply comparing their responses with those of other individuals.

A third converging source of evidence for decreasing fuzziness of category boundaries can be seen in the trend for children's categories to become narrower or more exclusive with age, in the sense that the number of emotion categories into which each

excerpt was placed decreased with age. Again, this finding is consistent with Bullock and Russell's (1986) research on the categorisation of emotional facial expressions.

As Alexander and Enns (1988) point out, the notion of age-related changes in the fuzziness of children's category boundaries makes sense of the apparently paradoxical literature on children's lexical generalisation, some of which argues that children typically overextend words, and some that they err by underextending words. A child with a fuzzy concept will *both* over- and under-extend that concept in comparison to individuals with a more clearly defined concept: on the one hand, they will overextend the concept to dissimilar exemplars; on the other, they will simultaneously be underextending the concept to similar exemplars. So, for instance, this study found a substantial proportion of 5-year-olds using 'sad' as a descriptor for excerpts classified by adults as expressing dissimilar emotions like happiness and excitement (overextension), and a correspondingly low proportion (compared with adults) using the term 'correctly'. Both of these tendencies decreased with age.

As Alexander and Enns (1988) comment, what this analysis suggests is that the concepts of younger children will share some features or properties with the concepts of older children and adults (accounting for the largely above-chance agreement with their categorisations), but that they will also contain some unique properties (accounting for overextension) and be missing some properties (accounting for underextension). Another way of looking at this is to consider a preschooler's prototype for a 'scared' piece of music, for instance, to share some features of the adult prototype, but also to exclude some features of the adult prototype and to include some features which are not part of the adult prototype.

### 7.4.1. Systematic inaccuracies

The idea that children's emotion concepts change slightly with age is consistent with the finding of this study that the majority of 5- and 7-year-olds chose an alternative emotional categorisation to that of adults for as many as a third of the excerpts

listened to, whereas 9-year-olds agreed with the adult emotion choice for 78% of the excerpts .

In addition, the argument that children's concepts for some emotions are similar, but not equivalent to, those of adults can help to explain why excerpts identified by adults as happy tended to be interpreted by a large proportion of children as excited, but not vice versa: when children disagreed with the adult categorisation of excited excerpts, their response was most likely to be scared. Viewed from the perspective of prototype theory, children tend to overextend the concept of excitement to include happy excerpts and to underextend it to excited excerpts, tending rather to label these as scared. This suggests that younger children's concept of excitement is something closer to happy than the adult meaning of the word. This interpretation is supported by the finding that for children (but not for adults) excited judgments were linked to consonant harmonies, which subjects of all ages associated with the expression of pleasant emotions. This interpretation is also consistent with previous research (Bullock & Russell, 1984, 1986; Manstead, 1993; Russell, 1990; Russell & Ridgeway, 1983; Trabasso, Stein & Johnson, 1981 in Stein & Jewett, 1984) suggesting that children tend to use the words 'excited' and 'surprised' to refer to positive events and facial expressions, whereas adults use them to refer to high arousal states which may be either positive or negative.[1]

Age-related changes in emotion concepts help to explain why the majority of 7-year-olds interpreted the happy excerpts as excited, while the majority of 5-year-olds interpreted the excited excerpts as scared. What they do not explain, however, is why systematic tendencies to confuse particular emotions - sadness and calmness, happiness and excitement, fear and anger as well as fear and excitement - were found for all age groups.

[1] It is also possible that this tendency may have been encouraged by the smiling 'excited' facial expression - it is conceivable that children may have placed more weight on the pictures than the adults, who might have concentrated more on the meaning of the emotion words.

One of these systematic 'errors' - the confusion of fear and anger - has also been found in previous studies (Cunningham & Sterling, 1988; Meerum Terwogt & van Grinsven, 1988, 1991). Meerum Terwogt and van Grinsven (1991) have attempted to explain this finding in terms of limitations in children's perspective-taking ability: children, it is argued, have difficulty differentiating the emotion evoked in themselves by an angry piece of music (fear) from the emotion expressed in the music.

Comments made by some of the 5-year-olds about excerpts which they classified as scared do lend some support to this hypothesis: "...like it's going to cut off your head" (excerpt 14); "sounds like a herd of elephants" (excerpt 2); "like a big monster stepping" (excerpt 10); here the judgment of fear presumably reflects the children's own response (the music is scary) rather than the emotion attributed to the monster, elephants, or aggressor. However, the finding that adults apparently often made the same 'mistake' suggests that it cannot be entirely attributed to children's cognitive egocentricity - particularly as recent research (Wellman et al., 1995) has suggested that even children as young as 2 can clearly distinguish between their own emotions and those of others, and accept that they can differ. Perhaps, then, it is more closely related to the music itself. In fact, several adults accounted for their difficulty in making the anger/fear distinction by explaining that some excerpts brought to mind a scene involving both an aggressor and a victim, a pursuer and a pursued. Their choice of emotion category therefore depended on which side they identified with - perhaps as a result of their personality or current mood. Support for this idea can be found in the open-ended emotional descriptions of the excerpts given by the pilot subjects: excerpt 14 ('The Fall of Saigon'), for example, evoked responses which included 'dread' as well as 'strength' and 'assertiveness'. One could hypothesise that because of children's relative vulnerability and powerlessness compared to adults, given such a scenario they would be more likely to identify with the weaker or 'scared' party, perhaps explaining their higher base rates for using the label 'scared'. From this perspective, anger and fear categorisations represent alternative ways of interpreting a musical excerpt, rather than a correct response and an error. Unlike Meerum Terwogt and van Grinsven's hypothesis, this account is able to explain why scared excerpts are sometimes interpreted as angry, and not only vice versa: musical expressions of fear

and anger are not polar opposites, as Dolgin and Adelson (1990) suggest, but rather contain a number of features in common, and overlap to a substantial degree.

The results of this investigation suggest that the same can be said of other pairs of emotions such as calmness and sadness, happiness and excitement, and fear and excitement. In particular, they provide support for the idea that emotion concepts are interrelated in a systematic fashion, with the degree of overlap between different emotions being explicable according to an underlying structure based on the dimensions of degree of pleasure and degree of activation (Russell, 1978, 1980, 1983, 1991b; Russell & Bullock, 1985, 1986; Russell & Ridgeway, 1983).

## 7.5. Dimensions

The results of this study provide three converging sources of evidence for the contention that the dimensions of pleasure and activation are a fundamental aspect of the interpretation of emotions in music for both adults and children - a suggestion that has been supported by several previous studies (Asada & Ohgushi, 1991; Meerum Terwogt & van Grinsven, 1991; Wedin, 1972).

Firstly, results on Task 1 showed that all three groups were able to distinguish between musical excerpts expressing similar degrees of pleasure while varying on the underlying dimension of degree of activation, and that they were most accurate at distinguishing the excerpts expressing sadness or calmness (low activation emotions, according to the structural model) from those expressing the other four emotions, which all fall into the high activation quadrants of the model. In addition, Task 2 showed that (with the exception of 5-year-olds' distinction between calm and sad excerpts) subjects of all ages were able to distinguish between excerpts expressing similar degrees of activation on the basis of the pleasantness or unpleasantness of the expressed emotions. This is consistent with Kratus's (1993) finding that 6- and 12-year-old children could distinguish between happy and sad music as well as between excited and calm music at well above chance levels.

A second source of evidence in favour of the dimensional approach to emotion concepts comes from the high correlations between the pilot subjects' pleasure and activation ratings of the musical excerpts on Mehrabian and Russell's (1974) semantic differential scales, and the emotion categories into which the excerpts were subsequently placed. Although this finding is perhaps not particularly surprising given that the semantic differential scales included distinctions such as happy-sad and excited-calm which were also used in the main study, they do support the idea that children and adults perceive emotions in music as varying in similarity in a systematic and quantitative way.

Additional support for an analysis of emotion concepts in terms of underlying dimensions comes from an examination of the 'errors' made by subjects: those choices that did not conform to the adult standard tended to fall into the pattern that would be predicted on the basis of the circumplex model. In other words, they were more likely to be selections of emotions expressing similar degrees of pleasure and activation than any other choice.

In fact, understanding subjects' emotion choices in terms of underlying dimensions can help to shed some light on the apparent discrepancies between this study and previous research (as well as between some of the previous studies themselves, and between the different tasks in this study) regarding the varying abilities of children to interpret particular emotions 'accurately'. From a dimensional perspective, children's performance on emotion detection tasks will vary according to the particular set of emotion labels with which they are provided as response alternatives: the closer these emotions fall to each other according to the structural model, the more difficult they will be to distinguish. The finding in this study that the relative consistency (that is, agreement with the adult modal choice) with which different emotions were identified in music varied from task to task is quite consistent with this idea: the 'accuracy' with which subjects identified musical expressions of particular emotions varied depending on the specific set of emotion categories with which they were provided as response alternatives. For example, calm and sad excerpts tended to be accurately identified more often than other excerpts on task 1, in which they had to be distinguished from

happy and excited, and angry and scared, excerpts respectively. However, they were identified correctly *less* often than all the other excerpts on task 2, which required subjects to differentiate calm and sad excerpts from each other. Similarly, angry and happy excerpts tended to be identified relatively well (compared to excerpts expressing other emotions) on task 2, which required that they be distinguished from each other. However, subjects were relatively poor at identifying excerpts expressing these states on tasks 1 and 3, where the response alternatives included emotions such as excitement and fear which fall much closer to happiness and anger respectively on the structural model. This finding supports Russell and Bullock's (1986) contention that children's superior performance for 'happy' stimuli in previous research might simply reflect the fact that in the majority of studies it is the only positive emotion included, and hence is maximally distinct from the other emotion categories provided (typically the three negative emotions of sadness, anger and fear). This might also help to explain why Kratus's (1993) 6-year-old subjects were no less consistent in their emotional categorisations of music than 12-year-olds: his subjects had only to choose between two emotions representing opposite poles of the pleasure and arousal dimensions, and may therefore have had an easier task than subjects who are presented with more closely overlapping emotions like anger and fear.

The finding of this study that 'errors' were most often characterised by similarity along the dimension of activity (for example, the confusion of calm and sad excerpts with each other, rather than with happy and angry excerpts respectively) and that subjects' emotion judgments were more highly correlated with arousal than pleasure ratings on Mehrabian and Russell's (1974) scales is also consistent with previous research suggesting that the activity dimension is more salient than the pleasure dimension in distinguishing auditory expressions of emotion[2] (Campbell, 1942; Dimitrovsky, 1964; Meerum Terwogt & van Grinsven, 1991; Scherer & Oshinsky, 1977), and with theoretical arguments such as those of Hanslick (1891/1986) and Langer (1942) which contend that it is the motion or dynamic aspects of emotion that

[2] A similar tendency can be observed in the responses of the pilot subjects, who tended to agree on whether excerpts expressed low or high activation emotions, but sometimes disagreed about whether high-arousal excerpts were excited (positive) or agitated (negative), or whether low-arousal excerpts were peaceful or sad.

are best expressed in music. This salience of the activation dimension in interpreting musical expressions of emotion appears to contrast with the interpretation of facial expressions, where the pleasure-displeasure dimension may be more important (Felleman et al., 1983), and confusion of emotions expressing similar degrees of pleasure but different degrees of activation (for instance anger and sadness) is more common.

Despite the overall support for the dimensional approach to emotion concepts, however, this study also revealed two minor trends in the pattern of responses which were not consistent with what the structural model would predict. The tendency of younger children to interpret sad and calm excerpts as scared more often than the model would lead us to expect is in accordance with Bullock and Russell's unexpected finding in their 1984 study that subjects tended to apply a label of 'afraid' to the sad facial expression more often than to the adjacent angry and disgusted facial expressions, as well as with Dolgin and Adelson's (1990) finding that children tended to confuse frightened- and sad-sounding musical excerpts. A possible explanation for this finding has been provided by Meerum Terwogt and van Grinsven (1991), who point out that while there are prototypical types of fear which involve high arousal and motivate a person to act (to escape from the threatening situation), there are also others which block every activity and are immobilising (being "petrified" or "paralysed with fear"). These two different types of fear could fit in at two different places in the structural model, and the latter type may have more in common with the low activation emotions of sadness and calmness than with the high activation emotions of anger and excitement, particularly as fear and sadness also share the common feature of being low in potency or dominance (in other words, the subject is in a weak position).

The tendency of younger children to interpret angry excerpts as calm more frequently than would be expected is more difficult to explain. As is the case with fear, it might be that there are different prototypical types of anger. One adult subject in this study pointed out that anger can refer both to "fury" and to "anger kept inside". Scherer and Oshinsky (1977) have made a similar distinction between "hot" and "cold" anger, and

argue that the two might be expressed by different combinations of acoustic cues. While it is possible, therefore, that 'cold' anger may be confused with low arousal emotions, it is not clear why 'calm' is chosen more often than 'sad', which has a closer resemblance to anger in terms of hedonic tone. Perhaps it is simply the case that some younger children are unsure about the meaning of the term 'calm', and might have heard it used by a parent or teacher in the context of anger ("Calm down!"). Alternatively, because the calm facial expression was the most neutral in terms of the degree of pleasure expressed, perhaps younger children tended to select this expression for excerpts whose hedonic tone was neither clearly positive nor negative.

These exceptions are relatively minor, and by no means invalidate the structural model. Nevertheless, they do suggest that emotion categories may be interrelated in a way slightly more complicated than that captured by Bullock and Russell's (1984, 1986) 2-dimensional model (perhaps incorporating other dimensions such as potency) and that a single emotion category might have different prototypical forms which take on slightly different positions in 2-dimensional space. A similar argument in favour of the idea that many emotions have several prototypes has been provided by Kövecses (1995).

## 7.6. Directions for future research

An important question which merits closer examination in future research concerns whether children are able to discriminate emotions in terms of broad dimensions *before* they are able to sort them into specific categories, or whether knowledge of categories and dimensions develops simultaneously, with children perhaps acquiring ideas about the degree of activation and pleasure implied by certain emotion concepts along with other features of the emotion script. Russell and Bullock's (1985, 1986; Russell, 1989) studies of children's interpretation of facial expressions have presented evidence that children as young as two or three can use dimensions to understand emotional expressions before they can categorise these expressions accurately. Since even the youngest subjects in this study were able to make quite reliable distinctions

between different emotion categories, we would have to study still younger children to determine whether their knowledge of emotion categories is preceded by knowledge of dimensions.

In addition, there are numerous other factors that might contribute to the emotional connotations of music and which must still be systematically examined if we are to refine our understanding of music's expression of emotion. These include the relationship between judgment and felt emotion: is it possible, for example, for adults and children to interpret a piece of music to represent extreme grief and yet be totally unmoved by it? If so, how? Other issues which require further study include the extent to which excerpts reflect more than one emotion, the effects of a person's current mood and of the order in which excerpts are heard, as well as of the social context of listening. In addition, although we have made some progress in this regard, a clear and full specification of the prototypical musical features associated with each emotion category still needs to be achieved - assuming that this daunting task is possible. Future research might also benefit from the inclusion of a structured set of probe questions designed to determine - in a more systematic way than was possible in this study - the basis of subjects' responses, although problems of children's limited vocabulary would need to be kept in mind. In addition, a wider variety of musical styles could be incorporated. Although the present study used more contemporary, popular and 'mundane' music than most previous investigations which have tended to concentrate on classical art music, numerous other musical styles including folk/ethnic music, rhythm and blues, heavy metal and so on deserve more specific focus in continuing investigations. In particular, it should be pointed out that although an effort was made in this study to select music that would be reasonably accessible to children, the excerpts still tended to be fairly complex and 'adult-centred' (Chris Wildman, personal communication, 12 December 1995). In other words, this study was investigating children's knowledge of the adult world and its music, rather than their interpretations of music specifically aimed at, or even created by, children. It is possible that the present results might had been different if simple traditional folk melodies or other, more 'child-centred' forms of music had been used. Finally, the question of what personal characteristics and experiences contribute to the ability to

interpret emotion in music also remains open, and future research might also attempt to address this issue.

In fact, the range of influences which might potentially contribute to the meaning of music for listeners of different ages is almost infinite, and much remains to be learned. Significant progress is, however, being made, and this study represents one more contribution towards understanding how it is that music is able to convey emotions in ways that can be recognised by listeners, and how the ability to interpret emotions in music develops with increasing age.

# CONCLUSIONS

In attempting to further our knowledge of how children's interpretation of emotion in music develops, this dissertation has presented theoretical arguments and experimental evidence which suggest that the conceptual network used by humans to understand emotion concepts involves fuzzy, prototypically organised categories related according to underlying dimensions of degree of pleasure and degree of activation or arousal. The application of this approach to the investigation of children's interpretation of emotion in music has provided some significant and meaningful findings that are not evident in results obtained in previous studies which have taken a more traditional approach to this issue, and simply evaluated children's responses as accurate or inaccurate against an adult standard.

The evidence presented here suggests that this traditional approach is inadequate for two main reasons. Firstly, the assumption that a particular emotional categorisation of a musical excerpt - whether it is the judgment of a musical expert, or of the majority of adults - is objectively correct, and that emotional development involves a straightforward registering of this universal truth, has been shown to be problematic. Rather, the notion of fuzzy categories implies that the emotional 'meaning' of musical excerpts is always more-or-less ambiguous for ordinary listeners, and that more than one emotion label may be justifiably applied to a single excerpt.

Secondly, it appears that simply evaluating children's performance as accurate or inaccurate against an adult norm may underestimate their ability to interpret emotional expressions in music: the evidence gathered here suggests that children's interpretation of emotion in music is quite systematic and meaningful, even though it is not exactly the same as that of adults. Children's so-called 'errors' are in fact not mistakes as such, but rather manifestations of their conceptual system. The major developmental change seen in this conceptual system is that as children grow older, so

their emotion categories become less fuzzy: that is, their emotion concepts become narrower, more specific and more exclusive, and the boundaries between different emotion concepts more stable.

Even preschool children, however, are clearly aware of at least some of the conventional emotional meanings expressed through music. To what extent these associations between music and emotion are culture-bound is uncertain: although there may well be a universal, biological basis for some emotional connotations in sounds, it is also clear that the average 5-year-old today has already learnt much about the emotional connotations of music from its dramatic use in films and television and its association with particular kinds of situations.

As Kratus (1993) points out, one implication of this finding for music teachers is that there is little need for them to teach primary school students to associate emotions with music, because they can already do this. On the other hand, descriptions like 'happy music' and 'sad music' could be used effectively as starting points in teaching children about musical elements, because unlike technical labels such as pitch or major and minor, these are labels that children readily understand.

The findings of this research also support the common-sense assumptions underlying the use of music to enhance (or occasionally reduce or ironically counter) the emotional qualities of words, actions or images in film and television dramas and commercials (Bruner, 1990; Radocy & Boyle, 1979), as well as to work diagnostically and therapeutically with people and their emotional problems (Behrens & Green, 1993). Although music therapists who work with children should be aware that the emotional connotations of music tend to be more ambiguous for children than for adults, and hence avoid making assumptions about how a particular selection will be interpreted, music might be a particularly useful tool for helping young children who find it difficult to talk about their emotions to explore and express them nonverbally.

What this dissertation does not intend to suggest, however, is that summing up or classifying musical excerpts with neat labels like happy, sad, or angry does justice

either to music as a means of emotional expression, or to those who appreciate it. The emotions conveyed by 'expressive' music are frequently complex, subtle, multifaceted and shifting; often they have a peculiar quality of their own which is only revealed by attending to the music itself, and which resists description in any words at all (Barwell, 1986). The fact that verbal labels are only a coarse approximation of what we perceive in music is a perennial problem when talking about it and the other arts, albeit one which might have to be accepted if we are to study the arts empirically. Nevertheless, we should continue to seek other, nonverbal methods which can be used as alternatives or supplements.

We also need to be careful of falling into the trap of believing that something is only fully understood when it can be expressed verbally: it may well be that music conveys the ambiguous, overlapping, fluid and dynamic nature of emotions *better* than words (Appleton, 1993; Langer, 1942; Mendelssohn, 1842 in Cooke, 1959; Osborne, 1982; Radocy & Boyle, 1979). After all, words, as we have seen, are merely labels for fuzzy categories: markers which help to make sense of, and allow us to communicate economically about, an almost infinite number of possible combinations of subjective feelings, cognitive appraisals, experiences, expressions, and so on, by focusing on their common features and indicating their most general character. And if music *can* express the nature of emotions better than words, this raises the question of whether it - and the other arts - might be able to stimulate children's perceptive, emotional and cognitive development by widening their understanding of the possibilities of human affect, and by helping them to make finer discriminations among their own feelings and those of others (Clynes, 1986; Gardner, 1983; Parsons, 1982; Reimer, 1989). Parsons (1982), for example, has put forward the idea that an individual's ability to empathise with others might be more finely guided and stimulated by a good artist than by the less carefully structured interactions of social life. And the musicologist John Blacking (1987, p.118) has argued that

> The development of the senses and the education of the emotions through the arts are not merely desirable options. They are essential both for balanced action and the effective use of the intellect.

Such questions about the significance and pedagogical uses of music go far beyond the issues with which this dissertation has been concerned, and it did not even attempt to answer them. Indeed, after centuries of debate about music's continuing importance in human experience, it is quite possible that a simple and universally acceptable explanation of music's connection with emotion will never be found. Nevertheless, through a combination of philosophical argument and empirical investigations that cut across the fields of psychology, linguistics, aesthetics, education, anthropology and of course also music itself, our knowledge of how we perceive and are affected by music gradually does seem to be expanding. This thesis represents one more step in that direction, carrying us a little further towards the goal of coming to a clear and complete understanding of how, during the course of development, we acquire knowledge about emotions, about music, and about the complex and fascinating relationship between the two. And in so doing, perhaps it will help to resolve a little of the confusion and controversy surrounding the scientific study of two of the most important, fundamental, and yet in many ways least understood aspects of what it means to be human.

# REFERENCES


Abeles, H.F. (1980). Responses to music. In D.A. Hodges (Ed.), Handbook of Music Psychology (pp.105-140). Lawrence, N.J.: National Association for Music Therapy.

Albas, D.C., McCluskey, K.W. & Albas, C.A. (1976). Perception of the emotional content of speech: A comparison of 2 Canadian groups. Journal of Cross-Cultural Psychology, 7(4), 481-490.

Alexander, T.M. & Enns, J.T. (1988). Age changes in the boundaries of fuzzy categories. Child Development, 59, 1372-1386.

Appleton, J. (1993). Epilogue: Implications for contemporary musical practice. In T. J. Tighe & W.J. Dowling (Eds.), Psychology and music: the understanding of melody and rhythm. Hillsdale, NJ: Laurence Erlbaum Associates.

Asada, M. & Ohgushi, K. (1991). Perceptual analyses of Ravel's Bolero. Music Perception, 8(3), 241-250.

Averill, J.R. (1986). The acquisition of emotions in adulthood. In R. Harré (Ed.), The social construction of emotions (pp. 98-118). Oxford: Blackwell.

Bartlett, J.C. & Dowling, W.J. (1980). Recognition of transposed melodies: A key-distance effect in developmental perspective. Journal of Experimental Psychology: Human Perception and Performance, 6(3), 501-515.

Barwell, I. (1986). How does art express emotion? The Journal of Aesthetics and Art Criticism, 45(2), 175-181.

Beckwith, R.T. (1991). The language of emotion, the emotions, and nominalist bootstrapping. In D. Frye & C. Moore (Eds.), Children's theories of mind: Mental states and social understanding (pp. 77-95). Hillsdale: Laurence Erlbaum Associates.

Behrens, G.A & Green, S.B. (1993). The ability to identify emotional content of solo improvisations performed vocally and on three different instruments. Psychology of Music, 21, 20-33.

Beldoch, M. (1964). Sensitivity to expression of emotional meaning in three modes of communication. In J.R. Davitz (Ed.), The communication of emotional meaning (pp. 13-29). New York: McGraw-Hill.

Berlyne, D.E. (1972). Affective aspects of aesthetic communication. In T. Alloway, L. Krames & P. Pliner (Eds.), Communication and affect: A comparative approach (pp. 97-118). New York: Academic Press.

Blacking, J. (1973). How musical is man? Seattle: University of Washington Press.

Blacking, J. (1987). 'A commonsense view of all music': Reflections on Percy Grainger's contribution to ethnomusicology and music education. Cambridge: Cambridge University Press.

Borke, H. (1971). Interpersonal perception of young children: Egocentrism or empathy? Developmental Psychology, 5(2), 263-269.

Borke, H. (1973). The development of empathy in Chinese and American children between three and six years of age. Developmental Psychology, 9(1), 102-108.

Bortolotti, S., D'Elia, P. & Whissell, C. (1993). When children talk about the causes of their emotions, how well do adults and other children understand which emotion they are talking about? Perceptual and Motor Skills, 77, 67-78.

Bretherton, I., Fritz, J., Zahn-Waxler, C. & Ridgeway, D. (1986). Learning to talk about emotions: A functionalist perspective. Child Development, 57, 529-548.

Brody, L.R. (1985). Gender differences in emotional development: A review of theories and research. Journal of Personality, 53(2), 102-149.

Brown, R. (1981). Music and language. In Documentary Report of the Ann Arbor Symposium: National Symposium on the Applications of Psychology to the Teaching and Learning of Music (pp. 233-264). Reston: Music Educators National Conference.

Bruner, G.C., II. (1990). Music, mood, and marketing. Journal of Marketing, 54, 94-104.

Budd, M. (1985). Music and the emotions: The philosophical theories. London: Routledge & Kegan Paul.

Budd, M. (1989). Music and the communication of emotion. The Journal of Aesthetics and Art Criticism, 47(2), 129-138.

Bullock, M. & Russell, J.A. (1984). Preschool children's interpretation of facial expressions of emotion. International Journal of Behavioral Development, 7, 193-214.

Bullock, M. & Russell, J.A. (1986). Concepts of emotion in developmental psychology. In C.E. Izard & P.B. Read (Eds.), Measuring emotions in infants and children (Vol. II, pp. 203-237). Cambridge: Cambridge University Press.

Burke, J. (1988). The illustrated dictionary of music. London: Sphere Books.

Bush, L.E. (1973). Individual differences multidimensional scaling of adjectives denoting feelings. Journal of Personality and Social Psychology, 25(1), 50-57.

Callen, D. (1982). The sentiment in musical sensibility. The Journal of Aesthetics and Art Criticism, 41(3), 381-393.

Campbell, I.G. (1942). Basal emotional patterns expressible in music. American Journal of Psychology, 55(1), 1-17.

Campos, J.J. & Barrett, K.C. (1984). Toward a new understanding of emotions and their development. In C.E. Izard, J. Kagan & R.B. Zajonc (Eds.), Emotions, Cognition, and Behavior (pp. 229-263). Cambridge: Cambridge University Press.

Camras, L.A. (1980). Children's understanding of facial expressions used during conflict encounters. Child Development, 51, 879-885.

Camras, L.A. & Allison, K. (1985). Children's understanding of emotional facial expressions and verbal labels. Journal of Nonverbal Behavior, 9(2), 84-94.

Clynes, M. (1986). Music beyond the score. Communication and Cognition, 19(2), 169-194.

Clynes, M. & Nettheim, N. (1982). The living quality of music: Neurobiological basis of communicating feeling. In M. Clynes (Ed.), Music, mind, and brain: The neuropsychology of music (pp. 47-82). New York: Plenum Press.

Cohen, A. J., Thorpe, L.A. & Trehub, S.E. (1987). Infants' perception of musical relations in short transposed tone sequences. Canadian Journal of Psychology, 41(1), 33-47.

Cooke, D. (1959). The language of music. London: Oxford University Press.

Costanzo, F. S., Markel, N.N. & Costanzo, P.R. (1969). Voice quality profile and perceived emotion. Journal of Counseling Psychology, 16, 267-270.

Crespo, E. (1986). A regional variation: Emotions in Spain. In R. Harré (Ed.), The social construction of emotions (pp. 209-217). Oxford: Blackwell.

Critchley, M. (1977). Ecstatic and synaesthetic experiences during musical perception. In M. Critchley & R.A. Henson (Eds.), Music and the brain. London: William Heinemann.

Crowder, R.G. (1984). Perception of the major/minor distinction: I. Historical and theoretical foundations. Psychomusicology, 4(1-2), 3-12.

Crowder, R.G. (1985). Perception of the major/minor distinction: III. Hedonic, musical, and affective discriminations. Bulletin of the Psychonomic Society, 23(4), 314-316.

Cunningham, J.G. & Sterling, R.S. (1988). Developmental change in the understanding of affective meaning in music. Motivation and Emotion, 12(4), 399-413.

Davidson, L., McKernon, P. & Gardner, H. (1981). The acquisition of song: A developmental approach. In Documentary Report of the Ann Arbor Symposium: National Symposium on the Applications of Psychology to the Teaching and Learning of Music (pp. 301-314). Reston: Music Educators National Conference.

Darwin, C. (1904). The expression of the emotions in man and animals (2nd ed.). London: John Murray.

Davies, S. (1991). [Review of Kivy, P. Sound Sentiment: an essay on the musical emotions]. The Journal of Aesthetics and Art Criticism, 49(3), 83.

Davitz, J.R. (1964a). A review of research concerned with facial and vocal expression of emotion. In J.R. Davitz (Ed.), The communication of emotional meaning (pp. 13-29). New York: McGraw-Hill.

Davitz, J.R. (1964b). Auditory correlates of vocal expressions of emotional meanings. In J.R. Davitz (Ed.), The communication of emotional meaning (pp. 101-112). New York: McGraw-Hill.

Davitz, J.R. (1969). The language of emotion. New York: Academic Press.

DePaulo, B.M. & Rosenthal, R. (1978). Age changes in nonverbal decoding as a function of increasing amounts of information. Journal of Experimental Child Psychology, 26, 280-287.

DePaulo, B.M. & Rosenthal, R. (1979). Age changes in nonverbal decoding skills: Evidence for increasing differentiation. Merrill-Palmer Quarterly, 25(2), 145-150.

Dimitrovsky, L. (1964). The ability to identify the emotional meaning of vocal expressions at successive age levels. In J.R. Davitz (Ed.), The communication of emotional meaning (pp. 69-86). New York: McGraw-Hill.

Dolgin, K.G. & Adelson, E.H. (1990). Age changes in the ability to interpret affect in sung and instrumentally-presented melodies. Psychology of Music, 18, 87-98.

Dowling, W.J. (1982). Melodic information processing and its development. In D. Deutsch (Ed.), The psychology of music (pp. 413-477). New York: Academic Press.

Dowling, W.J. & Harwood, D.L. (1986). Music in cognition. San Diego: Academic Press.

Dunn, J. (1995). Children as psychologists: The later correlates of individual differences in understanding of emotions and other minds. Cognition and Emotion, 9 (2/3), 187-201.

Edmonston, W.E., Jr. (1966). The use of the semantic differential technique in the aesthetic evaluation of musical excerpts. American Journal of Psychology, 79, 650-652.

Ekman, P. (1993). Facial expression and emotion. American Psychologist, 48(4), 384-392.

Ekman, P. & Friesen, W.V. (1971). Constants across cultures in the face and emotion. Journal of Personality and Social Psychology, 17(2), 124-129.

Ekman, P., Friesen, W.V., O'Sullivan, M., Diacoyanni-Tarlatzis, I., Krause, R., Pitcairn, T., Scherer, K., Chan, A., Heider, K., Le Compte, W.A., Ricci-Bitti, P.E., Masatoshi, T. & Tzavaras, A. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. Journal of Personality and Social Psychology, 53(4), 712-717.

Elson, L.C. (1905). Elson's music dictionary. Boston: Oliver Ditson.

Epstein, S. (1984). Controversial issues in emotion theory. In P. Shaver (Ed.), Review of Personality and Social Psychology (5): Emotions, relationships, and health (pp. 64-88). Beverly Hills, CA: Sage.

Farnsworth, P.R. (1969). The social psychology of music. Ames: Iowa State University Press.

Fehr, B. & Russell, J.A. (1984). Concept of emotion viewed from a prototype perspective. Journal of Experimental Psychology: General, 113(3), 464-486.

Fehr, B., Russell, J.A. & Ward, L.M. (1982). Prototypicality of emotions: A reaction time study. Bulletin of the Psychonomic Society, 20(5), 253-254.

Felleman, E.S., Barden, R.C., Carlson, C.R., Rosenberg, L. & Masters, J.C. (1983). Children's and adults' recognition of spontaneous and posed emotional expressions in young children. Developmental Psychology, 19(3), 405-413.

Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? Child Development, 60, 1497-1510.

Field, T.M. & Walden, T.A. (1982). Production and discrimination of facial expressions by preschool children. Child Development, 57, 136-142.

Francés, R. (1988). The perception of music (W.J. Dowling, Trans.) Hillsdale, NJ: Laurence Erlbaum Associates (Original work published 1958).

Francés, R. & Bruchon-Schweitzer, M. (1983). Musical expression and body expression. Perceptual and Motor Skills, 57, 587-595.

Gabriel, C. (1978). An experimental study of Deryck Cooke's theory of music and meaning. Psychology of Music, 6, 13-20.

Gardner, H. (1983). Frames of mind: The theory of multiple intelligences. New York: Basic Books.

Gaston, E.T. (Ed.) (1968). Music in therapy. New York: The Macmillan Company.

Gibbs, J.G. & Woll, S.B. (1985). Mechanisms used by young children in the making of empathic judgments. Journal of Personality, 53(4), 575-585.

Gordon, S.L. (1989). The socialization of children's emotions: Emotional culture, competence, and exposure. In C. Saarni & P.L. Harris (Eds.), Children's understanding of emotion (pp. 319-349). Cambridge: Cambridge University Press .

Griffiths, P.E. (1989). The degeneration of the cognitive theory of emotions. Philosophical Psychology, 2(3), 297-313.

Gross, A.L. & Ballif, B. (1991). Children's understanding of emotion from facial expressions and situations: A review. Developmental Review, 11, 368-398.

Gundlach, R.H. (1932). A quantitative analysis of Indian music. American Journal of Psychology, 44, 133-145.

Gundlach, R.H. (1935). Factors determining the characterization of musical phrases. The American Journal of Psychology, 47, 624-643.

Haack, P.A. (1980). The behavior of music listeners. In D.A. Hodges (Ed.), Handbook of music psychology (pp. 141-182). Lawrence, NJ: National Association for music therapy.

Hampton, P.J. (1945). The emotional element in music. Journal of General Psychology, 33, 237-250.

Hanslick, E. (1986). On the musically beautiful: A contribution towards the revision of the aesthetics of music (G. Payzant, Trans.) Indianapolis: Hackett (Original work published 1891).

Hargreaves, D.J. (1986). The developmental psychology of music. Cambridge: Cambridge University Press.

Harré, R. (1986). An outline of the social constructionist viewpoint. In R.Harré (Ed.), The social construction of emotions (pp. 2-14). Oxford: Blackwell.

Harris, P.L. (1989). Children and emotion: the development of psychological understanding. Oxford: Basil Blackwell.

Harris, P.L. & Saarni, C. (1989). Children's understanding of emotion: An introduction. In C. Saarni & P.L. Harris (Eds.), Children's understanding of emotion (pp. 3-24). Cambridge: Cambridge University Press.

Harwood, D.L. (1976). Universals in music: A perspective from cognitive psychology. Ethnomusicology, 20, 521-533.

Heinlein, C.P. (1928). The affective characters of the major and minor modes in music. Journal of Comparative Psychology, 8, 101-142.

Henkin, R.L. (1955). A factorial study of the components of music. Journal of Psychology, 39, 161-181.

Henkin, R.L. (1957). A reevaluation of a factorial study of the components of music. Journal of Psychology, 43, 301-306.

Hevner, K. (1935). The affective character of the major and minor modes in music. American Journal of Psychology, 47, 103-118.

Hevner, K. (1936). Experimental studies of the elements of expression in music. American Journal of Psychology, 48, 246-268.

Hevner, K. (1937). The affective value of pitch and tempo in music. American Journal of Psychology, 49, 621-630.

Izard, C.E. (1984). Emotion-cognition relationships and human development. In C.E. Izard, J. Kagan & R.B. Zajonc (Eds.), Emotions, Cognition, and Behavior (pp. 17-37). Cambridge: Cambridge University Press.

Izard, C.E. (1991). The psychology of emotions. New York: Plenum Press.

Johnson-Laird, P.N. & Oatley, K. (1989). The language of emotions: An analysis of a semantic field. Cognition and Emotion, 3(2), 81-123.

Kagan, J. (1984). The idea of emotion in human development. In C.E. Izard, J. Kagan & R.B. Zajonc (Eds.), Emotions, cognition, and behavior (pp. 38-72). Cambridge: Cambridge University Press.

Kastner, M.P. & Crowder, R.G. (1990). Perception of the major/minor distinction: IV. Emotional connotations in young children. Music Perception, 8(2), 189-202.

Kivy, P. (1989). Sound sentiment: An essay on the musical emotions including the complete text of *The Corded Shell*. Philadelphia: Temple University Press.

Kövecses, Z. (1995). Introduction: Language and emotion concepts. In J.A. Russell, J.M. Fernandez-Dols, A.S.R. Manstead & J.C. Wellenkamp (Eds.), Everyday conceptions of emotion: An introduction to the psychology, anthroplogy and linguistics of emotion (pp. 3-15). Dordrecht: Kluwer Academic Publishers.

Kratus, J. (1993). A developmental study of children's interpretation of emotion in music. Psychology of Music, 21, 3-19.

LaBarbera, J.D., Izard, C.E., Vietze, P. & Parisi, S.A. (1976). Four- and six-month-old infants' visual responses to joy, anger, and neutral expressions. Child Development, 47, 535-538.

Langer, Susanne K. (1942). Philosophy in a new key: A study in the symbolism of reason, rite, and art. New York: New American Library of World Literature.

Lewis, M. (1989). Cultural differences in children's knowledge of emotional scripts. In C. Saarni & P.L. Harris (Eds.), Children's understanding of emotion (pp. 350-373). Cambridge: Cambridge University Press.

Lieberman, L.R. & Walters, W.M., Jr. (1968). Effects of repeated listening on connotative meaning of serious music. Perceptual and Motor Skills, 26, 891-895.

Lutz, C. (1986). The domain of emotion words on Ifaluk. In R. Harré (Ed.), The social construction of emotions (pp. 267-288). Oxford: Blackwell.

Lutz, C. & White, G.M. (1986). The anthropology of emotions. Annual Review of Anthropology, 15, 405-436.

Lyons, W. (1992). An introduction to the philosophy of emotions. In Strongman, K.T. (Ed.), International review of studies on emotion (Vol II, pp. 295-313). Chichester: John Wiley & Sons.

Manstead, A.S.R. (1993). Children's representation of emotions. In Pratt, C. & Garton, A.F. (Eds.), Systems of representation in children: Development and use (pp. 185-210). Chichester: John Wiley & Sons.

Manstead, A.S.R. (1995). Children's understanding of emotion. In J.A. Russell, J.M. Fernandez-Dols, A.S.R. Manstead & J.C. Wellenkamp (Eds.), Everyday conceptions of emotion: An introduction to the psychology, anthropology and linguistics of emotion (pp. 315-331). Dordrecht: Kluwer Academic Publishers.

Manstead, A.S.R. & Edwards, R. (1992). Communicative aspects of children's emotional competence. In Strongman, K.T. (Ed.), International Review of Studies on Emotion, Vol. II. (pp. 167-196). Chichester: John Wiley & Sons.

Masters, J.C. & Carlson, C.R. (1984). Children's and adults' understanding of the causes and consequences of emotional states. In C.E. Izard, J. Kagan & R.B.Zajonc (Eds.), Emotions, cognition, and behavior (pp. 438-463). Cambridge: Cambridge University Press.

Matsumoto, D. (1992). American-Japanese cultural differences in the recognition of universal facial expressions. Journal of Cross-Cultural Psychology, 23(1), 72-84.

McCloskey, M.E. & Glucksberg, S. (1978). Natural categories: Well-defined or fuzzy sets? Memory and Cognition, 6(4), 462-472.

McCluskey, K.W., Albas, D.C., Niemi, R.R., Cuevas, C. & Ferrer, C.A. (1975). Cross-cultural differences in the perception of the emotional content of speech: A study of the development of sensitivity in Canadian and Mexican children. Developmental Psychology, 11(5), 551-555.

Meerum Terwogt, M. & Van Grinsven, F. (1988). Recognition of emotions in music by children and adults. Perceptual and Motor Skills, 67, 697-698.

Meerum Terwogt, M. & Van Grinsven, F. (1991). Musical expression of moodstates. Psychology of Music, 19, 99-109.

Mehrabian, A. & Russell, J.A. (1974). An approach to environmental psychology. Cambridge, MA: MIT Press.

Mesquita, B. & Frijda, N.H. (1992). Cultural variations in emotions: A review. Psychological Bulletin, 112(2), 179-204.

Moland, M. & Whissell, C.M. (1993). Children's understanding of facial expression of emotion: III. Adults' categorical and dimensional responses to children's drawings. Perceptual and Motor Skills, 77, 11-15.

Mull, H.K. (1949). A study of humor in music. American Journal of Psychology, 62, 560-566.

Nielzén, S. & Cesarec, Z. (1982). Emotional experience of music as a function of musical structure. Psychology of Music, 10(2), 7-17.

Parsons, M. (1982). Aesthetic development. In J.M. Broughton & D.J. Freeman-Moir (Eds.), The cognitive-developmental psychology of James Mark Baldwin: Current theory and research in genetic epistemology (pp. 389-433). Norwood: Ablex.

Peirce, C.S.S. (1955). Philosophical writings of Peirce (J. Buchler, Ed.). New York: Dover Publications.

Peters, S. (1990). A critical assessment of the development of musical concepts and skills with 5 yr old children. Early Child Development and Care, 54, 11-27.

Plutchik, R. (1962). The emotions: Facts, theories, and a new model. New York: Random House.

Profyt, L. & Whissell, C. (1991). Children's understanding of facial expression of emotion: I. Voluntary creation of emotion-faces. Perceptual and Motor Skills, 73, 199-202.

Radocy, R.E. & Boyle, J.D. (1979). Psychological foundations of musical behaviour. Springfield: Charles C. Thomas.

Reimer, B. (1989). A philosophy of music education (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Ridgeway, D., Waters, E. & Kuczaj, S.A., II. (1985). Acquisition of emotion-descriptive language: Receptive and productive vocabulary norms for ages 18 months to 6 years. Developmental Psychology, 21(5), 901-908.

Rigg, M.G. (1940). Speed as a determiner of musical mood. Journal of Experimental Psychology, 27, 566-571.

Rosch, E. (1975). Cognitive representations of semantic categories. Journal of Experimental Psychology: General, 104(3), 192-233.

Rosch, E. (1977). Human categorization. In N. Warren (Ed.), Studies in cross-cultural psychology (Vol. 1, pp. 3-49). London: Academic Press.

Rosenfeld, A. H. (1985). Music, the beautiful disturber. Psychology Today, 19(12), 48-56.

Russell, J.A. (1978). Evidence of convergent validity on the dimensions of affect. Journal of Personality and Social Psychology, 36(10), 1152-1168.

Russell, J.A. (1980). A circumplex model of affect. Journal of Personality and Social Psychology, 39(6), 1161-1178.

Russell, J.A. (1983). Pancultural aspects of the human conceptual organization of emotions. Journal of Personality and Social Psychology, 45(6), 1281-1288.

Russell, J.A. (1989). Culture, scripts, and children's understanding of emotion. In C. Saarni & P.L. Harris (Eds.), Children's understanding of emotion (pp. 293-318). Cambridge: Cambridge University Press.

Russell, J.A. (1990). The preschooler's understanding of the causes and consequences of emotion. Child Development, 61, 1872-1881.

Russell, J.A. (1991a). Culture and the categorization of emotions. Psychological Bulletin, 110(3), 426-450.

Russell, J.A. (1991b). In defense of a prototype approach to emotion concepts. Journal of Personality and Social Psychology, 60(1), 37-47.

Russell, J.A. & Bullock, M. (1985). Multidimensional scaling of emotional facial expressions: Similarity from preschoolers to adults. Journal of Personality and Social Psychology, 48(5), 1290-1298.

Russell, J.A. & Bullock, M. (1986). On the dimensions preschoolers use to interpret facial expressions of emotion. Developmental Psychology, 22(1), 97-102.

Russell, J.A., Lewicka, M. & Niit, T. (1989). A cross-cultural study of a circumplex model of affect. Journal of Personality and Social Psychology, 57(5), 848-856.

Russell, J.A. & Ridgeway, D. (1983). Dimensions underlying children's emotion concepts. Developmental Psychology, 19(6), 795-804.

Saarni, C. & Harris, P.L. (Eds.). (1989). Children's understanding of emotion. Cambridge: Cambridge University Press.

Scherer, K.R. (1979). Acoustic concomitants of emotional dimensions: Judging affect from synthesized tone sequences. In S.Weitz (Ed.), Nonverbal communication: Readings with commentary (2nd ed., pp. 249-253). New York: Oxford University Press.

Scherer, K.R. (1984). Emotion as a multicomponent process: A model and some cross-cultural data. In P. Shaver (Ed.), Review of personality and social psychology (5): Emotions, relationships, and health (pp. 37-63). Beverly Hills, CA: Sage.

Scherer, K.R. (1986). Vocal affect expression: A review and a model for future research. Psychological Bulletin, 99(2), 143-165.

Scherer, K.R., Banse, R., Wallbott, H.G. & Goldbeck, T. (1991). Vocal cues in emotion encoding and decoding. Motivation and Emotion, 15(2), 123-148.

Scherer, K.R. & Oshinsky, J.S. (1977). Cue utilization in emotion attribution from auditory stimuli. Motivation and Emotion, 1(4), 331-346.

Schlosberg, H. (1952). The description of facial expressions in terms of two dimensions. Journal of Experimental Psychology, 44(4), 229-237.

Schlosberg, H. (1954). Three dimensions of emotion. Psychological Review, 61(2), 81-88.

Scruton, R. (1993). Notes on the meaning of music. In Krausz, M. (Ed.), The interpretation of music: Philosophical essays (pp. 193-202). Oxford: Clarendon Press.

Shaver, P. (Ed.). (1984). Review of personality and social psychology (5): Emotions, relationships, and health. Beverly Hills, CA: Sage.

Shaver, P., Schwartz, J., Kirson, D. & O'Connor, C. (1987). Emotion knowledge: Further exploration of a prototype approach. Journal of Personality and Social Psychology, 52(6). 1061-1086.

Sloboda, J.A. (1985). The musical mind: The cognitive psychology of music. Oxford: Clarendon Press.

Sloboda, J.A. (1992). Empirical studies of emotional responses to music. In M.R. Jones & S. Holleran (Eds.), Cognitive bases of musical communication (pp. 33-46). Washington, DC: American Psychological Association.

Smiley, P. & Huttenlocher, J. (1989). Young children's acquisition of emotion concepts. In C. Saarni & P.L. Harris (Eds.), Children's understanding of emotion (pp. 27-94). Cambridge: Cambridge University Press.

Smith, K.D. & Tkel-Sbal, D. (1995). Prototype analyses of emotion terms in Palau, Micronesia. In J.A. Russell, J.M. Fernandez-Dols, A.S.R. Manstead & J.C. Wellenkamp (Eds.), Everyday conceptions of emotion: An introduction to the psychology, anthropology and linguistics of emotion (pp. 85-102). Dordrecht: Kluwer Academic Publishers.

Smith, S.T. & Smith, K.D. (1995). Turkish emotion concepts: A prototype approach. In J.A. Russell, J.M. Fernandez-Dols, A.S.R. Manstead & J.C. Wellenkamp (Eds.), Everyday conceptions of emotion: An introduction to the psychology, anthropology and linguistics of emotion (pp. 103-119). Dordrecht: Kluwer Academic Publishers.

Stecker, R. (1984). Expression of emotion in (some of) the arts. The Journal of Aesthetics and Art Criticism, 42(4), 409-418.

Stein, N.L. & Jewett, J.L. (1986). A conceptual analysis of the meaning of negative emotions: Implications for a theory of development. In C.E. Izard & P.B. Read (Eds.), Measuring emotions in infants and children (Vol.2, pp. 238-267). Cambridge: Cambridge University Press.

Stifter, C.A. & Fox, N.A. (1986). Preschool children's ability to identify and label emotions. Journal of Nonverbal Behavior, 10(4), 255-266.

Storm, C. & Storm, T. (1987). A taxonomic study of the vocabulary of emotions. Journal of Personality and Social Psychology, 53(4), 805-816.

Storr, A. (1992). Music and the mind. London: Harpercollins.

Sundberg, J. (1982). Speech, song, and emotions. In Clynes, M. (Ed.) Music, mind, and brain: The neuropsychology of music (pp. 137-149). New York: Plenum Press.

Thompson, W.F. & Robitaille, B. (1992). Can composers express emotions through music? Empirical Studies of the Arts, 10(1), 79-89 (From Psychological Abstracts, 1992, 79, Abstract no. 30566).

Tighe, T.J. & Dowling, W.J. (Eds.). (1993). Psychology and music: The understanding of melody and rhythm. Hillsdale, NJ: Laurence Erlbaum Associates.

Trainor, L.J. & Trehub, S.E. (1992). A comparison of infants' and adults' sensitivity to Western musical structure. Journal of Experimental Psychology: Human Perception and Performance, 18(2), 394-402.

Trehub, S.E. (1987). Infants' perception of musical patterns. Perception and Psychophysics, 41(6), 635-641.

Trehub, S.E. (1993). Music listening skills. In T.J. Tighe & W.J. Dowling (Eds.), Psychology and music: The understanding of melody and rhythm (pp. 161-176). Hillsdale, NJ: Laurence Erlbaum Associates.

Trehub, S.E., Bull, D. & Thorpe, L.A. (1984). Infants' perception of melodies: The role of melodic contour. Child Development, 55, 821-830.

Trehub, S.E., Thorpe, L.A. & Morrongiello, B.A. (1987). Organizational processes in infants' perception of auditory patterns. Child Development, 58, 741-749.

Tremblay, C., Kirouac, G. & Dore, F.Y. (1987). The recognition of adults' and children's facial expressions of emotions. Journal of Psychology, 121(4), 341-350.

Van Bezooijen, R., Otto, S.A. & Heenan, T.A. (1983). Recognition of vocal expressions of emotion: A three-nation study to identify universal characteristics. Journal of Cross-Cultural Psychology, 14(4), 387-406.

Walden, T.A. & Field, T.M. (1982). Discrimination of facial expressions by preschool children. Child Development, 53, 1312-1319.

Walker, A.S. (1982). Intermodal perception of expressive behaviors by human infants. Journal of Experimental Child Psychology, 33, 514-535.

Wedin, L. (1972). A multidimensional study of perceptual-emotional qualities in music. Scandinavian Journal of Psychology, 13, 241-257.

Wellman, H.M., Harris, P.L., Banerjee, M. & Sinclair, A. (1995). Early understanding of emotion: Evidence from natural language. Cognition and Emotion, 9(2/3), 117-149.

Whissell, C.M. & D'Elia, P. (1993). Adults' free-form identification of emotions from children's descriptions of their antecedents: A quantitative analysis. Perceptual and Motor Skills, 77, 3-9.

Wierzbicka, A. (1986). Human emotions: Universal or culture-specific? American Anthropologist, 88, 584-594.

Wierzbicka, A. (1992a). Defining emotion concepts. Cognitive Science, 16, 539-581.

Wierzbicka, A. (1992b). Talking about emotions: Semantics, culture, and cognition. Cognition and Emotion, 6(3/4), 285-319.

Williams, C.E. & Stevens, K.N. (1972). Emotions and speech: Some acoustical correlates. In S.Weitz (Ed.), Nonverbal communication: Readings with commentary (pp. 233-253). New York: Oxford University Press.

Young-Browne, G., Rosenfeld, H.M. & Horowitz, F.D. (1977). Infant discrimination of facial expressions. Child Development, 48, 555-562.

## Appendix 1: Semantic differential scales for hedonic tone and arousal (from Mehrabian & Russell, 1974)

### Instructions

This study is concerned with how people perceive emotions expressed in music. You are going to hear 16 short musical excerpts. I would like you to listen to each excerpt carefully, and then to rate it on the set of scales which you will find overleaf. In other words, you will rate the first musical excerpt on all the scales in set 1, the second musical excerpt on all the scales in set 2, and so on.

Each of the scales contains a pair of adjectives separated by seven spaces. Some of the pairs might seem unusual or irrelevant to the music, but you will probably feel that one adjective describes the music better than the other. So, for each pair, you should put a cross close to the adjective which you feel describes the music better. The more appropriate the adjective seems, the closer you put your cross to it. You should place your cross in the middle space only if you consider the music to be neutral on that scale, or if the scale is completely irrelevant.

Note:

1. Place your crosses in the middle of spaces, not on the boundaries.
2. Please mark the scales in the order they appear.
3. Be sure you mark every scale for every piece of music - do not leave any out.
4. Do not put more than one cross on a single scale.

Please make each item a separate and independent judgment. Do not look back and forth through the items. You should work at fairly high speed, because it is your first impressions I want. On the other hand, please do not be careless, because I want your true impressions.

## Scales utilised

### Hedonic tone

| | | |
|---|---|---|
| contented | ...\|...\|...\|...\|...\|...\|... | melancholic |
| relaxed | ...\|...\|...\|...\|...\|...\|... | bored |
| satisfied | ...\|...\|...\|...\|...\|...\|... | unsatisfied |
| pleased | ...\|...\|...\|...\|...\|...\|... | annoyed |
| happy | ...\|...\|...\|...\|...\|...\|... | unhappy |
| hopeful | ...\|...\|...\|...\|...\|...\|... | despairing |

### Arousal

| | | |
|---|---|---|
| jittery | ...\|...\|...\|...\|...\|...\|... | dull |
| wide awake | ...\|...\|...\|...\|...\|...\|... | sleepy |
| frenzied | ...\|...\|...\|...\|...\|...\|... | sluggish |
| aroused | ...\|...\|...\|...\|...\|...\|... | unaroused |
| excited | ...\|...\|...\|...\|...\|...\|... | calm |
| stimulated | ...\|...\|...\|...\|...\|...\|... | relaxed |

Following Mehrabian & Russell (1974), these 12 scales were combined and presented to subjects in random order, with half the scales reversed.

## Appendix 2: Verbal expressions of emotion

**CALM:** "It will be so peaceful. We shall listen to the birds and have a goat, and grow the best vegetables in Trentino."
Adapted from Ibbotson, E. (1985). A Company of Swans (p. 56). New York: Avon Books.

**HAPPY:** "Oh Danny. Didn't we have a glorious time! We did it, Danny. We pulled it off. That was the greatest time I've ever had in my whole life."
Adapted from Dahl, R. (1975). Danny The Champion of the World pp. 141, 145). Harmondsworth: Penguin.

**EXCITED:** "I wish tomorrow would come quickly. Suppose there is something really thrilling in that ruined building after all! Now I shan't be able to go to sleep tonight - I shall keep on and on thinking about tomorrow!"
Adapted from Blyton, E. (1976). The Treasure Hunters (p. 131). London: William Collins Sons & Co.

**SCARED:** "That man over there....Don't let him see me! Oh, can't we turn back, please...please....He's come to take me back!"
Adapted from Ibbotson, E. (1985). A Company of Swans (p. 171).

**ANGRY:** "How dare you, Harriet? How dare you argue with me! You will leave the table immediately and go to your room!"
Adapted from Ibbotson, E. (1985). A Company of Swans (pp. 23-24).

**SAD:** "What I can't figure out is why the good things always end. I don't want to go. I don't want this to be over. I want to stay, Torey." Adapted from Hayden, T. L. (1980). One Child (p. 206). New York: Avon Books.

## Appendix 3: Instructions to subjects (Task 1)

Hello, my name is Lauren and I'm here today because I want to find out what children know about feelings. I'd like you to help me do this by playing some special games that I've brought with me. In the first game, we're going to listen to some music that has a feeling. Sometimes when we listen to music, music sounds happy, and sometimes it sounds calm or excited. Today I'm going to play you a tape of some music, and when you listen your job will be to decide whether the music sounds happy or calm or excited. Look at your answer sheet. There are nine boxes. In each box there is a happy face, a calm face, and an excited face (show example). Can you point to the happy face in box 1? Can you point to the calm face in box 9? Can you point to the excited face in box 4? (Make sure that all the children are able to identify the faces in the proper boxes). I'm going to play you nine short pieces of music on the tape recorder. After you listen to each one, I want you to draw a circle round the face that you think best matches the feeling in the music. So if you think the first piece of music sounds happy, circle the happy face in box 1. If you think the music sounds calm, circle the calm face. If you think the music sounds excited, circle the excited face. Don't circle more than one face in the same box, and don't leave any boxes blank. If you change your mind about an answer, just cross out your old answer. This game has no right or wrong answers. I just want to know what each one of you thinks, so please don't look at anybody else's answers. Are there any questions?

...This is the first piece of music. Listen to it carefully.... Now draw a circle around the face in box 1 that you think best matches the feeling in the music....

## Appendix 4: distribution of emotion choices by age and task (n = 30)

| | | *Task 1* | | | | | | *Task 2* | | | | | | *Task 3* | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ex-cerpt | Age | Calm | Happy | Excited | Scared | Angry | Sad | Calm | Sad | Happy | Angry | Excited | Scared | Calm | Happy | Excited | Scared | Angry | Sad |
| 1 | 5yrs | 3 | 12 | 15 | | | | | | | | 16 | 14 | 3 | 4 | 8 | 9 | 4 | 2 |
| | 7yrs | 0 | 6 | 24 | | | | | | | | 24 | 5 | 1 | 2 | 25 | 1 | 0 | 0 |
| | 9yrs | 0 | 1 | 29 | | | | | | | | 19 | 11 | 1 | 1 | 22 | 6 | 0 | 0 |
| | Adult | 0 | 1 | 29 | | | | | | | | 18 | 12 | 2 | 0 | 22 | 2 | 4 | 0 |
| 2 | 5yrs | 10 | 10 | 10 | | | | | | | | 17 | 13 | 5 | 3 | 6 | 10 | 3 | 3 |
| | 7yrs | 12 | 8 | 10 | | | | | | | | 22 | 7 | 4 | 10 | 8 | 3 | 4 | 0 |
| | 9yrs | 7 | 1 | 22 | | | | | | | | 26 | 4 | 1 | 14 | 13 | 1 | 1 | 0 |
| | Adult | 2 | 5 | 23 | | | | | | | | 29 | 1 | 0 | 19 | 10 | 0 | 1 | 0 |

| | *Task 1* | | | | | | | *Task 2* | | | | | | *Task 3* | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ex-cerpt | Age | Calm | Happy | Excited | Scared | Angry | Sad | Calm | Sad | Happy | Angry | Excited | Scared | Calm | Happy | Excited | Scared | Angry | Sad |
| 3 | 5yrs | | | | 4 | 4 | 22 | 16 | 14 | | | | | 5 | 0 | 1 | 3 | 2 | 19 |
| | 7yrs | | | | 1 | 2 | 27 | 5 | 24 | | | | | 7 | 1 | 0 | 0 | 0 | 21 |
| | 9yrs | | | | 0 | 0 | 30 | 7 | 23 | | | | | 3 | 1 | 0 | 0 | 0 | 26 |
| | Adult | | | | 1 | 0 | 29 | 5 | 25 | | | | | 5 | 0 | 0 | 0 | 0 | 25 |
| 4 | 5yrs | 17 | 7 | 6 | | | | 13 | 17 | | | | | 15 | 1 | 2 | 1 | 1 | 10 |
| | 7yrs | 27 | 3 | 0 | | | | 8 | 21 | | | | | 11 | 6 | 2 | 2 | 0 | 8 |
| | 9yrs | 28 | 2 | 0 | | | | 10 | 20 | | | | | 25 | 0 | 0 | 1 | 0 | 4 |
| | Adult | 29 | 1 | 0 | | | | 14 | 16 | | | | | 24 | | | | | 6 |
| 5 | 5yrs | 5 | 16 | 9 | | | | | | 27 | 3 | | | 4 | 10 | 6 | 6 | 3 | 1 |
| | 7yrs | 5 | 15 | 10 | | | | | | 29 | 0 | | | 3 | 13 | 0 | 13 | 0 | 0 |
| | 9yrs | 9 | 21 | 0 | | | | | | 30 | 0 | | | 0 | 21 | 7 | 0 | 0 | 2 |
| | Adult | 4 | 23 | 3 | | | | | | 29 | 1 | | | 0 | 23 | 6 | 0 | 1 | 0 |
| 6 | 5yrs | 17 | 7 | 6 | | | | 19 | 11 | | | | | 12 | 6 | 2 | 1 | 5 | 4 |
| | 7yrs | 26 | 4 | 0 | | | | 26 | 3 | | | | | 21 | 3 | 2 | 0 | 0 | 3 |
| | 9yrs | 27 | 3 | 0 | | | | 29 | 1 | | | | | 26 | 2 | 0 | 0 | 0 | 1 |
| | Adult | 29 | 1 | 0 | | | | 29 | 1 | | | | | 20 | 7 | 0 | 0 | 0 | 3 |

| | *Task 1* | | | | | | | *Task 2* | | | | | | *Task 3* | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ex-cerpt | Age | Calm | Happy | Excited | Scared | Angry | Sad | Calm | Sad | Happy | Angry | Excited | Scared | Calm | Happy | Excited | Scared | Angry | Sad |
| 7 | 5yrs | | | | 11 | 16 | 3 | | | 4 | 26 | | | 3 | 0 | 3 | 7 | 15 | 2 |
| | 7yrs | | | | 21 | 6 | 2 | | | 6 | 23 | | | 1 | 0 | 2 | 13 | 13 | 0 |
| | 9yrs | | | | 10 | 20 | 0 | | | 2 | 28 | | | 0 | 0 | 3 | 6 | 20 | 0 |
| | Adult | | | | 11 | 19 | 0 | | | 0 | 30 | | | 0 | 0 | 1 | 8 | 21 | 0 |
| 8 | 5yrs | | | | 15 | 6 | 9 | 17 | 13 | | | | | 11 | 5 | 1 | 4 | 3 | 6 |
| | 7yrs | | | | 2 | 1 | 27 | 28 | 1 | | | | | 16 | 1 | 1 | 1 | 0 | 10 |
| | 9yrs | | | | 1 | 0 | 29 | 25 | 5 | | | | | 20 | 1 | 1 | 1 | 0 | 7 |
| | Adult | | | | 1 | 0 | 29 | 25 | 5 | | | | | 26 | 1 | 0 | 1 | 0 | 2 |
| 9 | 5yrs | 5 | 6 | 19 | | | | | | 23 | 7 | | | 2 | 13 | 7 | 2 | 3 | 3 |
| | 7yrs | 0 | 10 | 20 | | | | | | 27 | 2 | | | 0 | 8 | 20 | 0 | 1 | 0 |
| | 9yrs | 0 | 22 | 8 | | | | | | 30 | 0 | | | 0 | 14 | 16 | 0 | 0 | 0 |
| | Adult | 0 | 20 | 10 | | | | | | 29 | 1 | | | 0 | 19 | 10 | 0 | 1 | 0 |
| 10 | 5yrs | | | | 14 | 12 | 4 | | | | | 6 | 24 | 1 | 1 | 2 | 19 | 5 | 2 |
| | 7yrs | | | | 15 | 12 | 2 | | | | | 4 | 25 | 0 | 0 | 2 | 19 | 8 | 0 |
| | 9yrs | | | | 23 | 7 | 0 | | | | | 0 | 30 | 0 | 0 | 0 | 29 | 1 | 0 |
| | Adult | | | | 28 | 1 | 1 | | | | | 1 | 29 | 0 | 0 | 2 | 28 | 0 | 0 |

| | *Task 1* | | | | | | | *Task 2* | | | | | | *Task 3* | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ex-cerpt | Age | Calm | Happy | Excited | Scared | Angry | Sad | Calm | Sad | Happy | Angry | Excited | Scared | Calm | Happy | Excited | Scared | Angry | Sad |
| 11 | 5yrs | | | | 5 | 5 | 20 | 13 | 17 | | | | | 6 | 2 | 2 | 2 | 2 | 16 |
| | 7yrs | | | | 2 | 0 | 28 | 20 | 8 | | | | | 15 | 1 | 0 | 2 | 0 | 11 |
| | 9yrs | | | | 0 | 0 | 30 | 13 | 17 | | | | | 10 | 2 | 0 | 0 | 0 | 18 |
| | Adult | | | | 0 | 0 | 30 | 3 | 27 | | | | | 6 | 0 | 0 | 0 | 0 | 24 |
| 12 | 5yrs | | | | 6 | 5 | 19 | | | 5 | 25 | | | 8 | 4 | 4 | 5 | 3 | 6 |
| | 7yrs | | | | 3 | 1 | 26 | | | 6 | 23 | | | 13 | 0 | 1 | 4 | 0 | 11 |
| | 9yrs | | | | 1 | 0 | 29 | | | 6 | 24 | | | 11 | 0 | 0 | 3 | 0 | 16 |
| | Adult | | | | 1 | 1 | 28 | | | 21 | 9 | | | 13 | 4 | 1 | 1 | 0 | 11 |
| 13 | 5yrs | 7 | 14 | 9 | | | | | | 29 | 1 | | | 2 | 16 | 10 | 0 | 1 | 1 |
| | 7yrs | 3 | 7 | 19 | | | | | | 29 | 0 | | | 0 | 7 | 21 | 1 | 0 | 0 |
| | 9yrs | 0 | 13 | 17 | | | | | | 30 | 0 | | | 0 | 20 | 10 | 0 | 0 | 0 |
| | Adult | 0 | 25 | 5 | | | | | | 30 | 0 | | | 0 | 20 | 10 | 0 | 0 | 0 |

| | *Task 1* | | | | | | | *Task 2* | | | | | | *Task 3* | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ex-cerpt | Age | Calm | Happy | Excited | Scared | Angry | Sad | Calm | Sad | Happy | Angry | Excited | Scared | Calm | Happy | Excited | Scared | Angry | Sad |
| 14 | 5yrs | | | | 8 | 20 | 2 | | | | | 12 | 18 | 0 | 1 | 4 | 12 | 13 | 0 |
| | 7yrs | | | | 11 | 18 | 1 | | | | | 19 | 10 | 0 | 1 | 1 | 14 | 13 | 0 |
| | 9yrs | | | | 13 | 17 | 0 | | | | | 7 | 23 | 0 | 0 | 0 | 5 | 25 | 0 |
| | Adult | | | | 4 | 25 | 1 | | | | | 4 | 26 | 0 | 0 | 1 | 6 | 23 | 0 |
| 15 | 5yrs | | | | 6 | 18 | 6 | | | 16 | 14 | | | 7 | 4 | 0 | 11 | 6 | 2 |
| | 7yrs | | | | 12 | 17 | 1 | | | 9 | 20 | | | 2 | 1 | 6 | 7 | 13 | 0 |
| | 9yrs | | | | 3 | 27 | 0 | | | 1 | 29 | | | 0 | 1 | 5 | 23 | 1 | 0 |
| | Adult | | | | 3 | 26 | 0 | | | 2 | 28 | | | 0 | 1 | 6 | 10 | 13 | 0 |
| 16 | 5yrs | 5 | 12 | 13 | | | | | | | | 21 | 9 | 5 | 6 | 11 | 2 | 2 | 4 |
| | 7yrs | 1 | 27 | 2 | | | | | | | | 25 | 4 | 5 | 7 | 10 | 2 | 4 | 1 |
| | 9yrs | 1 | 27 | 2 | | | | | | | | 30 | 0 | 3 | 17 | 10 | 0 | 0 | 0 |
| | Adult | 0 | 29 | 1 | | | | | | | | 30 | 0 | 0 | 20 | 10 | 0 | 0 | 0 |
| 17 | 5yrs | | | | 17 | 7 | 6 | | | | | 18 | 12 | 4 | 2 | 7 | 9 | 5 | 3 |
| | 7yrs | | | | 21 | 7 | 2 | | | | | 10 | 19 | 2 | 3 | 2 | 15 | 7 | 0 |
| | 9yrs | | | | 18 | 11 | 1 | | | | | 23 | 7 | 1 | 1 | 7 | 14 | 7 | 0 |
| | Adult | | | | 22 | 8 | 0 | | | | | 26 | 4 | 0 | 2 | 16 | 8 | 4 | 0 |

| | *Task 1* | | | | | | | *Task 2* | | | | | | *Task 3* | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ex-cerpt | Age | Calm | Happy | Excited | Scared | Angry | Sad | Calm | Sad | Happy | Angry | Excited | Scared | Calm | Happy | Excited | Scared | Angry | Sad |
| 18 | 5yrs | 22 | 5 | 3 | | | | 13 | 17 | | | | | 11 | 5 | 0 | 2 | 0 | 12 |
| | 7yrs | 27 | 2 | 1 | | | | 2 | 26 | | | | | 15 | 1 | 0 | 1 | 0 | 12 |
| | 9yrs | 30 | 0 | 0 | | | | 9 | 21 | | | | | 17 | 0 | 0 | 0 | 0 | 13 |
| | Adult | 30 | 0 | 0 | | | | 15 | 15 | | | | | 19 | 1 | 0 | 0 | 0 | 10 |

## Appendix 5: Summaries of ANOVAs for accuracy scores (comparison with adult mode) on the three music tasks

**Task 1**

| Source of variation | SS | df | MS | F | p |
|---|---|---|---|---|---|
| Between-subjects | | | | | |
| Age (A) | 13.444 | 3 | 4.481 | 36.429 | < 0.000* |
| Error-between | 14.270 | 116 | 0.123 | | |
| Within-subjects | | | | | |
| Emotion (E) | 7.899 | 5 | 1.580 | 27.753 | < 0.000* |
| A × E | 2.959 | 15 | 0.197 | 3.465 | < 0.000* |
| Error-within | 33.016 | 580 | 0.057 | | |

* Indicates significant at $\alpha = 0.05$

**Task 2**

| Source of variation | SS | df | MS | F | p |
|---|---|---|---|---|---|
| Between-subjects | | | | | |
| Age (A) | 5.378 | 3 | 1.793 | 34.350 | < 0.000* |
| Error-between | 6.002 | 115 | 0.052 | | |
| Within-subjects | | | | | |
| Emotion (E) | 3.451 | 5 | 0.690 | 18.386 | < 0.000* |
| A × E | 0.842 | 15 | 0.056 | 1.494 | 0.102 |
| Error-within | 21.588 | 575 | 0.038 | | |

* Indicates significant at $\alpha = 0.05$

**Task 3**

| Source of variation | SS | df | MS | F | p |
|---|---|---|---|---|---|
| Between-subjects | | | | | |
| Age (A) | 10.627 | 3 | 3.542 | 28.040 | < 0.000* |
| Error-between | 14.528 | 115 | 0.126 | | |
| Within-subjects | | | | | |
| Emotion (E) | 10.246 | 5 | 2.049 | 23.696 | < 0.000* |
| A × E | 1.446 | 15 | 0.096 | 1.114 | 0.339 |
| Error-within | 49.723 | 575 | 0.086 | | |

* Indicates significant at $\alpha = 0.05$

## Appendix 6: Correlations between musical elements and emotional categorisations for each age group separately (task 3)

| EMOTION | MUSICAL ELEMENTS | AGE 5yrs | 7yrs | 9yrs | Adult |
|---|---|---|---|---|---|
| CALM | Tempo | -0.59* | -0.76* | -0.65* | -0.68* |
| | Rhythm | -0.75* | -0.79* | -0.78* | -0.82* |
| | Rhythmic activity | -0.53* | -0.66* | -0.54* | -0.60* |
| | Articulation | -0.53* | -0.66* | -0.50* | -0.49* |
| | Harmony | -0.38 | -0.39 | -0.36 | -0.34 |
| | Modality | 0.01 | -0.17 | 0.12 | 0.10 |
| HAPPY | Tempo | 0.27 | 0.38 | 0.43 | 0.41 |
| | Rhythm | 0.25 | 0.25 | 0.37 | 0.36 |
| | Rhythmic activity | 0.22 | 0.28 | 0.31 | 0.30 |
| | Articulation | 0.39 | 0.27 | 0.38 | 0.37 |
| | Harmony | -0.59* | -0.63* | -0.57* | -0.59* |
| | Modality | 0.47 | 0.65* | 0.51 | 0.54* |
| EXCITED | Tempo | 0.71* | 0.69* | 0.81* | 0.85* |
| | Rhythm | 0.59* | 0.56* | 0.63* | 0.64* |
| | Rhythmic activity | 0.61* | 0.61* | 0.73* | 0.80* |
| | Articulation | 0.55* | 0.46 | 0.45 | 0.54* |
| | Harmony | -0.19 | -0.36 | -0.27 | -0.00 |
| | Modality | 0.55 | 0.52 | 0.49 | 0.42 |

| | | AGE | | | |
|---|---|---|---|---|---|
| EMOTION | MUSICAL ELEMENTS | 5yrs | 7yrs | 9yrs | Adult |
| SCARED | Tempo | 0.40 | 0.27 | 0.36 | 0.30 |
| | Rhythm | 0.44 | 0.40 | 0.32 | 0.27 |
| | Rhythmic activity | 0.43 | 0.32 | 0.41 | 0.33 |
| | Articulation | 0.40 | 0.44 | 0.40 | 0.39 |
| | Harmony | 0.68* | 0.85* | 0.71* | 0.74* |
| | Modality | -0.18 | -0.51 | -0.42 | -0.42 |
| ANGRY | Tempo | 0.12 | 0.29 | 0.02 | 0.09 |
| | Rhythm | 0.53* | 0.59* | 0.43 | 0.53* |
| | Rhythmic activity | 0.21 | 0.35 | 0.09 | 0.16 |
| | Articulation | 0.28 | 0.43 | 0.24 | 0.25 |
| | Harmony | 0.51* | 0.69* | 0.53* | 0.51* |
| | Modality | -0.16 | -0.29 | -0.18 | -0.33 |
| SAD | Tempo | -0.67* | -0.80* | -0.75* | -0.69* |
| | Rhythm | -0.79* | -0.87* | -0.73* | -0.67* |
| | Rhythmic activity | -0.71* | -0.83* | -0.80* | -0.72* |
| | Articulation | -0.83* | -0.81* | -0.82* | -0.82* |
| | Harmony | -0.11 | -0.10 | -0.05 | -0.04 |
| | Modality | -0.59* | -0.60* | -0.74* | -0.73* |
| PLEASANT | Tempo | 0.17 | 0.20 | 0.24 | 0.25 |
| | Rhythm | -0.00 | 0.00 | -0.01 | -0.03 |
| | Rhythmic activity | 0.12 | 0.16 | 0.20 | 0.20 |
| | Articulation | 0.21 | 0.02 | 0.12 | 0.19 |
| | Harmony | -0.72* | -0.85* | -0.84* | -0.76* |
| | Modality | 0.72* | 0.84* | 0.95* | 0.93* |

| EMOTION | MUSICAL ELEMENTS | AGE 5yrs | 7yrs | 9yrs | Adult |
|---|---|---|---|---|---|
| ACTIVE | Tempo | 0.74* | 0.86* | 0.87* | 0.86* |
| | Rhythm | 0.91* | 0.92* | 0.94* | 0.94* |
| | Rhythmic activity | 0.74* | 0.82* | 0.83* | 0.82* |
| | Articulation | 0.82* | 0.81* | 0.81* | 0.81* |
| | Harmony | 0.27 | 0.28 | 0.27 | 0.26 |
| | Modality | 0.43 | 0.44 | 0.37 | 0.40 |

* Indicates significant at $\alpha = 0.05$

## Appendix 7: Glossary of musical terms

CONSONANCE: An accord of sounds agreeable and satisfying to the ear

DISSONANCE: The sounding together of notes which produce a discord or disturb the ear

HARMONICS: The secondary sounds resonating (often barely audibly) at fixed intervals above the fundamental or principle tone

HARMONY: A combination of notes sounded simultaneously to produce a chord

LARGO: Slow

LEGATO: Notes are connected in a close, smooth, graceful manner

MAJOR SCALE/MODE: An eight-note scale consisting of a succession of whole-tone gaps except for a semitone between the third and fourth degrees and between the seventh and eighth degrees; that mode or scale in which the interval between the first and third notes comprises two whole tones or steps (See also **mode, scale, semitone, tone**)

MELODY: A succession of notes of varying pitch and rhythm which carry the 'tune' or dominant theme

MINOR SCALE/MODE: As the major scale, but with a semitone introduced between the second and third degrees and between the sixth and seventh degrees, with slight variations in certain formations; that scale or mode in which the interval between the first and third notes consists of one-and-a-half tones (See also **mode, scale, semitone, tone**)

METER: The pattern of successive rhythmic pulses produced by notes of similar or varying length within a bar

MODE: The arrangement of notes in a scale; the scale used to develop harmony, e.g. major, minor, doric, plagal (See also **scale, harmony**)

OVERTONES: See HARMONICS

PITCH: The height or depth of a sound, determined by the frequency or rate of vibration of the voice or instrument

PRESTO: Fast

RHYTHM: The systematic grouping of musical notes within a pattern which establishes their relationships with one another in time

SCALE: A ladder of ascending or descending successive notes

SEMITONE: A half-sized step between two notes, as between C and C sharp

STACCATO: Notes are detached, distinct, separated from each other

TEMPO: The pace or speed at which a piece of music is played; the rapidity at which the natural rhythmic accents follow each other

TIMBRE: Tone colour; the quality of the sound produced by different voices or instruments

TONE: An interval of a major second; the whole step between two notes, as between C and D

[Information drawn from Burke's (1988) and Elson's (1905) music dictionaries and definitions provided by Peters (1990)].