Emotion and the Experience of Listening to Music

A Framework for Empirical Research

Matthew Montague Lavy

Jesus College, Cambridge

This dissertation is submitted for the degree of Doctor of Philosophy

April 2001

Contents

Declaration

I declare that this thesis consists entirely of my own work, except where explicitly stated to the contrary. This thesis fits within the prescribed word limit of 80,000 words, excluding the Appendices (which consist largely of computer code).

Acknowledgements

I would like to thank the following people whose help has been invaluable during the researching and writing of this thesis: Dr Ian Cross (who has given me advice and support throughout my time at Cambridge); colleagues at the Jesus College IT Department; my parents and family; friends, particularly Nicola Phillips and Alistair Turnbull for their suggestions and numerous insightful comments. This research was funded by a grant from the Arts and Humanities Research Board, and my parents.

Abstract

That music evokes emotion is a well-known and uncontested fact. Rather more contentious have been the numerous attempts by philosophers, writers and musicians over the centuries to explain the phenomenon. In recent years, the development of cognitive psychology has led to renewed interest in the field; a growing number of music psychologists are devoting their energies to the empirical examination of various aspects of musically evoked emotion. Despite the wealth of data fast amassing, however, there exist few theoretical accounts of emotional response to music written from a music-psychological perspective within which empirical studies can be understood and upon which they can build. Furthermore, accounts that do exist have traditionally made rigid distinctions between *intrinsic* and *extrinsic* sources of emotion, distinctions that do not fit well with our understanding of emotional antecedents in other domains.

This thesis presents the foundations of a model of emotional response to music that places the experience of listening to music squarely within the wider frame of human engagement with the environment. Instead of presenting a categorization of music or an analysis of cultural or individual semantic tokens, the model develops four basic assumptions concerning listeners and their relationship to music:

1. *Music is heard as sound.* The constant monitoring of auditory stimuli does not suddenly switch off when people listen to music; just like any other stimulus in the auditory environment, music exists to be monitored and analyzed.

2. *Music is heard as human utterance.* Humans have a remarkable ability to communicate and detect emotion in the contours and timbres of vocal utterances; this ability is not suddenly lost during a musical listening experience.

3. *Music is heard in context.* Listeners do not exist in a vacuum: music is always heard within the context of a complex web of knowledge, thoughts and environment, all of which can potentially contribute to an emotional experience.

4. *Music is heard as narrative.* Listening to music involves the integration of sounds, utterances and context into dynamic, coherent experience. Such integration, far from being a phenomenon specific to music listening, is underpinned by generic narrative processes.

The first part of the thesis introduces the four components of the model, reviews existing empirical and theoretical research that supports its premises, and considers its ramifications. The discussion reveals that despite an abundance of evidence pointing to the importance of narrative for affective responses to music, virtually no empirical work has addressed the issue directly. Hence, the second part of the thesis presents three experiments that constitute a preliminary attempt to do so. First is an experiment that investigates the interaction between music and listening context in the evocation of emotional response. It presents participants with musical excerpts in conjunction with explicit extra-musical narratives in order to demonstrate how readily music *binds* with extra-musical context to form a dynamic, coherent whole. The second and third experiments seek to demonstrate that such binding is not specific to music but is an example of the workings of more generic cognitive processes that underpin narrative comprehension. In addition, both of these experiments are intended to exemplify research paradigms that could be used in future empirical research on the narrative processing of music and its role in the evocation of emotion.

The thesis argues that the Sound-Utterance-Context-Narrative model constitutes a good framework for empirical work because it is specific enough to provoke detailed research questions and methodologies, but generic enough that a theoretically complete answer to all the questions it poses would constitute a comprehensive understanding of emotional response to music. Its over-arching claim is that an understanding of emotional response to music can only be attained by the development of models that refrain from treating music as a privileged class of object with intrinsic emotional properties, and instead consider the act of listening to music as a perfectly ordinary human activity.

Part I

Accounting for Musically Evoked Emotion

Chapter 1

The need for a model

Philosophers, writers and musicians have been investigating the power of music to evoke emotions for millennia; the pursuit is probably as old as the phenomenon of music itself. Over the last few decades the development of music psychology—with its promise of scientific objectivity—has sparked a renewed flurry of activity in this field. Literally thousands of papers have been published that present data relating to various aspects of phenomena involving music and emotion from many disparate disciplines. For the music psychologist, however, there has been a single fundamental question: what is it about music that enables it to evoke emotions in listeners?[1]

We know a considerable amount about the effects of music both from systematic observation of human behaviour in the real world and from the results of controlled laboratory experiments. An abundance of evidence indicates not only that people regularly use music deliberately as a mood induction tool (e.g. Sloboda, 2000), but also that listening to music in everyday circumstances can lead to moderated mood and altered cognition (e.g. Bruner, 1990). For evidence of the centrality of music to the emotional lives of many Americans, at least, we need look no further than Frey (1985), whose massive survey reported that eight percent of all crying episodes in the USA were evoked either directly or indirectly by music.

[1]The terms "emotion", "emotions", "emotional response" and "affective response" are used informally throughout this thesis to refer to a broad category of pychological and physiological states that would, in common parlance, be described as emotional. "Emotional response to music", then, refers to any affective response whether momentary or longer-lived (mood-like) that has been induced by music. In using emotion terms somewhat interchangeably, this thesis in no way wishes to denigrate the substantial work that has sought to differentiate between various affective phenomena, but merely to avoid making distinctions that are still largely a matter of contention even within the field of emotion psychology itself (see e.g. Davidson, 1994). As the thesis develops, it will become apparent that this decision has no impact on the model presented. There are some points in later chapters where it has been helpful to distinguish between categories of affect (for want of a better word); in these cases, both the distinctions and their rationale has been made explicit.

Sometimes, music evokes intense feelings of sadness and joy, which can occasionally lead even to strong physical reactions or overt physical behaviour, such as shouting, screaming or crying (Gabrielsson and Lindström, 2000); less extreme physical reactions, such as piloerection and lumps in the throat, are commonly cited musically induced phenomena (e.g. Sloboda, 1991). Emotional responses to music are correlated with physiological functioning, as evidenced by measurable effects on the sympathetic and parasympathetic systems (e.g. Iwanaga and Tsukamoto, 1997; Krumhansl, 1997b) and by detectable changes in neural activity (e.g. Gerra, Zaimovic, Franchini, Palladino, Guicastro, Reali, Maestri, Caccavari, Delsignore, and Brambilla, 1998; Panksepp, 1997). As a complement to these findings, clinical observation and the success of music therapy (see Bunt, 1994) in helping patients with psycho-physiological complaints ranging from autism (e.g. Keats, 1995) to dementia (e.g. Pinkney, 1999) also stand testament to music's power to affect psychological and physiological state.

In the light of these observations, it is hardly surprising that a central quest amongst music psychologists interested in emotion has been the isolation of those factors in music that might be responsible for the many observed effects. An armoury of robust and complementary experiments has indicated beyond a shadow of doubt that the broad emotional characterisation of music as perceived by listeners familiar with western musical idioms is mediated by parameters such as mode, intensity, timbre, and tempo. The most important parameter in determining musical expression appears to be mode: major keys are consistently associated with positive emotions, whereas minor ones are associated with negative emotions (e.g. Hevner, 1936; Scherer and Oshinsky, 1997); the ability to discriminate the valence of emotional expression on the basis of mode has even been noted in three-year-old children (Kastner and Crowder, 1990). Tempo has also been found to be highly influential: fast music is associated with joy and activity, whereas slow music is associated with sadness or solemnity; tempo interacts with other parameters in the generation of perceived emotional expression, such as with mode, melodic contour or, most notably intensity (e.g. Rigg, 1940). Intensity alone has not been shown to be unambiguously associated with any specific emotional expression, but has a mediating effect on tempo: a high intensity makes fast music seem more energetic than its lower intensity counterpart whereas slow music appears more serious or solemn; a low intensity, on the other hand, makes slow music appear sad-der, more contemplative and gloomy than its louder counterparts (e.g. Wedin, 1972). The effect of timbre is somewhat more complex and cannot easily be summarised (see Chapters Two and Three), but one relatively unambiguous finding has been that timbral quality of a musical sound consistently affects perception of emotional tension (e.g. Nielzen and Olsson, 1993).

Studies that have investigated the influence of musical parameters on perceived emotional expression have gathered and analysed data using a variety of techniques, some of which have involved measuring physiological correlates of emotion in the hope of finding evidence for a direct relationship between the state of a particular musical parameter and emotion evocation. Most, however, have relied on introspection, asking listeners to rate musical extracts on semantic differential scales, to match extracts with predefined lists of adjectives, or sometimes even to give free reports of responses. Some studies have used real music in their experiments, others have used specially designed stimuli, and yet others have used hybrid techniques involving systematic manipulation of real musical examples. Each of these approaches is beset by methodological problems of varying degrees of severity, as has been discussed several times elsewhere (see e.g. Rigg, 1964; Dowling and Harwood, 1986; Sloboda, 1996); a perpetually recurring problem concerns the difficulty of demonstrating that a listener's ability to perceive emotional expression in music necessarily equates with emotion evocation (an issue to which we shall return shortly). Despite these partially unresolved issues, research on broad emotional characterisations of musical parameters is succouring an impressive corpus of extremely useful data.

All the research mentioned so far has been concerned with general emotional characterisations of music, looking at those parameters that tend to remain constant throughout a piece (e.g. music is slow, or is in the minor key or has a bright timbre). Perhaps the single most important characteristic of music, however, is its dynamic quality—the fact that it constantly changes over time; certainly this is crucial to the entire edifice of western tonal musical structure, which is built on a dynamic flux of tensions and resolutions both in the melodic and harmonic domains (see e.g. Narmour, 1990, 1992; Schmuckler, 1989). Concomitantly, although music moderates mood, and although characterisations of static attributes of a musical work may be at least partly responsible for this—the difficulty of equating perception of expression and emotion evocation notwithstanding—emotional response to music can be intense and *momentary*. In recent years, much research has concentrated on dynamic aspects of musical structure that may be responsible for the evocation of emotions, and here too, consistent patterns of results have emerged: a growing body of evidence indicates that certain small-scale musical structures, such as appoggiaturas and suspensions, evoke emotional responses reliably in accultured listeners (e.g. Sloboda, 1991; Waterman, 1996, 1997) and that listeners reliably detect and respond to changing levels of tension that emerge as a piece of music unfolds (e.g. Krumhansl, 1996). Perception of melodic and harmonic structure is not the only source of musical dynamics to have been the subject of extensive research: another concerns timing deviations in performances, where perceived emotionality of a performance has consistently found to be correlated with deviations from strictly accurate timing patterns (Repp, 1997; Sloboda, 1997). As serious research in this area is a relatively new phenomenon, there are few established research methods; investigators in this field have often encountered difficulties attempting to capture listeners' momentary responses without disturbing the listening experience. Developments in computer technology made a huge difference here, and have effectively provided a *de facto* research paradigm in which listeners are asked to "track" perceived tension—or some other dimension—in real time. Results take the form of fluctuating continuous data in one or two dimensions derived from the position of a joystick or a computer mouse, or the pressure applied to a pressure-sensitive device (e.g. Madsen and Fredrickson, 1993; Schubert, 1998). Stimuli tend to be real musical extracts, often manipulated in some way.

The extensive data fast amassing concerning both perceived emotional characterisations in music and emotion evocation as a result of the dynamics of musical structure constitute a formidable canon. All work that attempts to isolate musical parameters—static or dynamic—that may be responsible for the evocation of emotional responses, however, is confounded by an apparently inescapable paradox: however strong the evidence for agreement between listeners on the evocative properties of certain musical parameters, there exists an equally powerful body of evidence to suggest that two accultured listeners can respond to the same piece of music in vastly different ways; even worse, the same listener can respond to the same piece of music in different ways at different times. As Sloboda (1992, p.38) put it, one can "listen to the same recording on two difference occasions and be moved to tears on one of them, while remaining completely detached on the other". The implication for research on broad emotional characterisation of music is worrying: perception of emotional expression encoded within musical parameters certainly does not inexorably lead to the evocation of those emotions in a listener. As for musical dynamics, it would appear that listeners may be moved by musical moments, but not under all circumstances! In other words, musical antecedents of emotion are not universally definable. All this research may tell us a lot about the general consensus of accultured listeners, which perhaps in turn tells us much about musical convention, but it does not necessarily tell us anything about musical evocation of emotions.

While music psychologists have devoted their efforts to revealing what aspects of music might be responsible for the evocation of emotional reactions, many researchers in fields more closely allied to social psychology have considered the extra-musical factors that may impinge on emotional responses to music. Results, mostly based on analysis of listener introspections, have demonstrated that the extent to which people are moved even by the most paradigmatically emotionally expressive music can depend on their mood (Sloboda, 1992), the situation in which the music is heard (Konečni, 1982), or even the motivation for listening (North and Hargreaves, 2000). In addition, listeners are frequently moved by a piece of music if it reminds them of previous periods in their lives, whether by bringing to mind specific episodes (e.g. Baumgartner, 1992) or general eras (e.g. Schulkind, Hennis, and Rubin, 1999), regardless of the musical content. In short, when emotional responses are attributed to music, it is often not the music at all that causes the response but something completely different which happens to be associated with the music in the listener's mind. These observations, of course, tell us nothing about the interaction between emotional characterisations, musical dynamics and extra-musical "contaminants"—an issue that surprisingly little work has addressed—but they do serve as a useful complement to the plethora of music-based research and should serve to put many of the findings into perspective.[2]

The difficulty of reconciling what has been discovered about the evocative properties of music with what is known about the idiosyncrasies of listening is symptomatic of a more general problem that besets almost all research in the field, namely that despite the wealth of data, there exist few theoretical accounts of emotional response to music within which

[2]Much of the research mentioned in paragraphs above will be considered in greater detail later on in this thesis. A forthcoming book by Sloboda and Juslin (in press) promises to provide a series of thorough literature reviews covering the full gamut of research topics relating to music and emotion from a number of perspectives.

empirical studies can be understood and upon which they can build. To be sure, several psychologists, philosophers and musicologists have presented theories that impinge on the subject, but very few have written from a music psychological perspective. As a result, experimental work remains phenomenological and atheoretic: whilst we know a great deal about a number of isolated phenomena, we have very little understanding about how these phenomena interact within a wider frame. Although some empirical researchers have grounded their work in terms of more general theories of emotion, hardly any attempt has been made to weave the various facets of emotionality in music into a single coherent picture that can constitute the needed "wider frame" upon which future work can be based.[3] Music psychologists seem to have been particularly reluctant to embrace findings concerning the influence of extra-musical phenomena when considering emotional response to music.

Of the theoretical accounts of emotional response to music that do exist, by far the most thorough—and most oft cited—is that of Meyer (1956), which presents a detailed analysis of the structures that underlie much of western tonal music and considers how those structures conspire to evoke emotional reactions in listeners. Meyer's great contribution was lucidly to describe the dynamics of western tonal music as building a series of tensions and resolutions, and to demonstrate how listening to such music involves the constant generation of expectancies on many levels, which are confirmed or violated as the music unfolds. Meyer grounded his largely musicologically-inspired thesis in the context of a theory of inhibition-based arousal,[4] arguing that expectancy violations in music are particularly prone to evoke emotional reactions because "emotion or affect is aroused when a tendency to respond is arrested or inhibited"(p.14) and unlike in the real world, where unresolved tendencies are "dissipated in the press of irrelevant events ... in art inhibition of tendency becomes meaningful because the relationship between the tendency and its resolution is made explicit and apparent"(p.23). This conception of music and the listening process fits well with arousal-based cognitive models of emotions, and a variety of empirical studies have shown it to be rather robust. The vast number of empirical studies based on or inspired by Meyer's work even today stand testament to the great debt owed to him by music psychologists studying emotion.

[3]A similar criticism has previously been levelled specifically at research related to emotional cues in musical performances: "researchers interested in music performance have been reluctant to turn to the psychology of emotion for theoretical guidance"(Juslin, 1997, p.412).

[4]He cites MacCurdy (1925), who sees arousal as resulting from the repression of instinctive behaviour.

Whilst being an extremely important work for music psychologists, Meyer's theory has one serious (self-declared) limitation: it only concerns itself with emotional responses to structural characteristics of western tonal music. As it is not generalisable enough to embrace other musical genres, or to explain responses that are not directly related to the dynamics of musical structure, it remains an insightful discourse on what is likely to be a significant source of emotional response to music for a listener steeped in the western musical tradition, but is too narrow in scope to form the basis of a general framework for empirical research.[5]

A theory by Dowling and Harwood (1986) represents a more generic approach, which whilst not nearly as detailed as that of Meyer—it is not intended to be—does offer a wide overall perspective on what it is to which listeners might be responding when they respond emotionally to music. Drawing both on the work of the philosopher Charles Peirce (1935) and on a review of the core empirical literature, Dowling and Harwood model emotion in music as existing on three levels, that of *index, icon* and *symbol*. Music acting as index evokes responses by provoking associations in a listener's mind between that music and something extra-musical. Indexical association can be sparked as a deliberate compositional ploy, such as in the use of the French and Russian national anthems in Tchaikovsky's *1812 Overture*, or the taped fire sirens in *Poème electronique* by Varèse; alternatively, they may belong exclusively in the domain of a particular listener, for whom a particular piece of music might spark related memories.[6] Either way, music acting as index works in a Pavlovian (1927) way, and can evoke emotion to the extent that the associand constitutes an emotional stimulus for the listener.

The iconic level represents the sounds and patterns of a musical surface, or the ebb and flow of a musical line. Drawing heavily on Langer (1951, 1953), Dowling and Harwood suggest that "music mimics the form ... of emotional life"(p.206), in support of which they present a thorough review of the empirical research on emotion characterisations in music (covering similar ground to the empirical work reviewed in this chapter). If the concept of iconic ebb and flow seems somewhat amorphous, that is allegedly because "the representation of emotion in musical icons is necessarily vague"(p.207).

[5]To be fair to Meyer, he did not present his theory as a framework for empirical research. In fact, he virtually rejected the possibility that the subject could usefully be studied using empirical psychological methods.

[6]Dowling and Harwood provide by way of illustration an excerpt from Proust (1934) in which recognition of a violin sonata sparked memories of a love affair from years before.

In contrast to icons, symbols are emotional by virtue of their place and function within a musical structure: Mandler's (1984) theory of autonomic arousal is cited as the basis for a conceptualisation of western music much akin to that of Meyer (1956). The symbolic level of representation, therefore, is only available to an accultured listener, who has an understanding of the musical idioms in question. Between them, index, icon and symbol are intended to cover the whole gamut of musical features to which listeners might respond emotionally.

Dowling and Harwood are by no means alone in their tripartite categorisation; the distinctions they draw between index, icon and symbol reflect broadly the implicit distinctions made by many empirical researchers in the field, because they map well on to phenomenologically identified categories of emotional response to music. Even the brief taxonomy of empirical work in the field given above reveals three distinct branches: one contains studies of the way in which listeners associate musical works with extra-musical concepts (indexical associations); another, of the various parameters that affect "emotionality" of a musical surface (icons); yet another, of musical structure and its ability to evoke emotional responses in the accultured listener (symbols). The distinction between the second and third branches may not always be entirely clear—empirical work inspired by Meyer sits somewhere in the middle as it concerns both the (iconic) flow of a musical line and the dynamics of musical (symbolic) structure—but the first branch is almost entirely segregated from the rest of the tree, not least because it is the branch of least interest to many music psychologists.

However closely the distinction between index, icon and symbol maps on to empirical research, it does not fare so well when considered in the light of the phenomenon of emotional response itself. Not only does the difficulty of placing Meyer's work and the empirical evidence that supports it show the distinction between icon and symbol to be artificial at best, but also music theory and plain common sense conspire to indicate that symbol and index can often be interchangeable. For example, a perfect cadence, which is both a highly recognisable and conventional musical unit, is certainly part of a symbol structure—in itself encapsulating a complete expectancy generation and resolution cycle—but also has the potential to act indexically because as a tokenistic, recognisable, closed unit, it could well spark associations both cultural and idiosyncratic for a listener;[7] it seems plausible, in fact, that in many circumstances, emotional response to musical structures such as perfect

[7] This latter point has been made ludicly by Cooke (1959) and will be considered further at several points later on in the thesis.

cadences could be understood much better as a phenomenon allied to indexical association than to one allied to the tensions and resolutions of a musical dynamic. In short, the distinctions are merely an artifact of theorising; they may broadly reflect the landscape of phenomenologically-inspired empirical research, but do not well represent any essential reality.

An artificial distinction between index, icon and symbol is really the product of a more fundamental distinction made by many current conceptualisations of emotional response to music, namely the distinction between sources of emotion that are ascribed to music itself, so-called *intrinsic* sources, and those sources that are only tangentially related to music, so-called *extrinsic* sources. It is inevitable that aspects of a research area as wide in scope as emotional response to music will attract researchers working within different, possibly even incompatible paradigms, and it is certainly of no cause for alarm that researchers whose agenda is primarily social will be interested in music's place in the emotional lives of listeners, whereas those whose interest lies squarely within the realm of music will concentrate more on the emotion evoking properties of musical parameters. The real problem is that for the music psychologist, the distinction is not an epistemological but an ontological one: it is not assumed merely to be different paradigmatic agenda that cause some researchers to concentrate on extrinsic factors while others study intrinsic factors; rather, underlying the dichotomy seen in the empirical and theoretical work has been the assumption, often explicitly asserted, that intrinsic and extrinsic sources of emotion are ontologically distinct entities. A corrollory is the assumption that emotional responses evoked by extrinsic sources and those evoked by intrinsic sources are quite different phenomena. A central tenet of this thesis is that as long as such an assumption remains, a coherent model of emotional response to music will always be unattainable and empirical research will continue to be phenomenological and atheoretic.

Music psychologists investigating intrinsic sources of emotional response make heavy use of music theoretic concepts. Such concepts, which have evolved over several hundred years, provide countless pre-built units and dissection routines that can be used to categorise and control individual musical parameters in the hope of isolating those that encode emotional signals. As has already been noted, the approach has for the most part been rather successful. Unfortunately, this success has led to a strange insularity amongst many researchers, perpetuated by the implicit belief that there exists an absolute distinction between intrinsic and extrinsic sources of emotion: musicologically inspired research need not be exasperated in its failure to account for idiosyncratic response differences that depend on listeners and situations, because these can be considered to be caused by extrinsic factors and are therefore irrelevant to the pursuit of discovering how intrinsic factors evoke emotions. As Dowling and Harwood (1986, p.205) put it: "powerful though these conditioned reactions [to extrinsic sources of emotion] can be ... there is nothing particularly *musical* about the process."

A listener's reaction to extrinsic sources of emotion may not be exclusively musical, but they are intrinsic to the listening process; surely it is extrinsic factors such as cultural context and previous experience that help give so-called intrinsic factors their evocative power. It seems that through all the categorisation, one aspect of emotional response to music that has been largely forgotten is the listener! For the majority of researchers, the fundamental question has inevitably been "what are the features of music that evoke emotional responses?". By contrast, very few have asked the question "what are the features of listeners that make them interpret music as an emotional stimulus?". It is hardly surprising that, with a few exceptions, work in this field appears oblivious to more general emotion theories: what is the need for an emotion theory when the object under scrutiny is music and not the listener?

If musicologically inspired research is ever simultaneously to form a satisfactory understanding of why certain musical parameters and the dynamics of musical structure appear to be sources of emotion, to resolve the difficulty of relating perception of musical expression to evocation of emotion, to address listener idiosyncracy and to leave phenomenology behind, it will have to be situated within a model of emotional response to music at the heart of which is placed the listener. Such a model must avoid arbitrary categorisations based on *a priori* concepts and musicological or philosophical assumptions; it must not see response to music as a phenomenon that is somehow special and isolated from other facets of human emotion.[8] Instead it must seek to situate what we know about emotional response to music squarely within the frame of everyday human experience.

It is here claimed that far from being an unattainable goal, the plethora of empirical research that has been carried out in the field over the last few decades, together with an ever-increasing body of apposite theory and experiment from related research areas, does in

[8] Perhaps Meyer's one disservice to music psychology was to see musical emotions as somehow different from those of everyday life.

fact suggest coherent and consistent patterns upon which a model could be based. Providing the foundations of such a model will be the purpose of this thesis.

Instead of attempting a categorisation of music itself, the model will be founded upon four basic premises concerning listeners and their relationship to music:

1. *Music is heard as sound.* In daily life, humans—in common will all other hearing animals—are constantly monitoring the auditory environment, listening for signs of danger, scanning for anything that might require action. For the most part, such monitoring happens non-consciously, but when something untoward is heard, attention immediately focuses on that stimulus; meanwhile, the body readies itself for action and the hearer experiences a sense of arousal. The model's first assumption is that this constant monitoring process does not suddenly switch off when people listen to music; just like any other stimulus in the auditory environment, music exists to be monitored and analysed.

2. *Music is heard as human utterance.* A theory propounded by many contemporary evolution theorists states that there is one type of sound to which humans are particularly sensitive, namely, the sound of the human voice. An abundance of empirical data lends credence to the hypothesis that we have a remarkable ability to communicate and detect emotion in the contours and timbres of vocal utterances; it is assumed that this ability is active and important during the process of listening to music.

3. *Music is heard in context.* Listeners do not exist in a vacuum: music is heard within the context of a complex web of knowledge, thoughts and environment, all of which can potentially contribute to an emotional experience.

4. *Music is heard as narrative.* Many psychologists, anthropologists and sociologists have noted that central to the human ability to make sense of all facets of the world around us is an ability to construct narratives. These narratives can take many forms, but their function is always essentially the same: to link disparate fragments of information and to find connections and commonalities between different stimuli. Narratives, like the situations and stimuli from which they are constructed, may well contain information of great relevance to the well-being of their constructor; therefore, they are often emotionally valenced. The present model assumes that making sense of music, just like

making sense of other stimuli and structures, involves the construction of narratives.

The remainder of Part One of this thesis will lay out the case for viewing emotional response to music as response to sound, utterance, context and narrative by attempting to illustrate how many of the known phenomena relating to music and emotion fit into a wider, more general frame. It will hope to demonstrate how a conceptualisation of emotional response to music in terms of sound, utterance, context and narrative can ultimately lead to a model that is empirically based, simple and, ideally, predictive, and that such a model could provide a much needed framework for empirical research. It will argue that disregarding all theoretical distinctions between so-called intrinsic and extrinsic sources of emotional response will lead not only to extant empirical data being seen in a coherent, consistent light, but could also pave the way to new, hitherto unexplored avenues of enquiry that could help further elucidate the nature of one of life's enigmatic mysteries, namely emotional response to music.

Chapter 2

Music as Sound

"Although most psychological work treats appraisal as a high-level cognitive process ... it is clear from studies of animals and people that stimuli are first evaluated at a lower (unconscious) level prior to, and perhaps independent of, higher-level appraisal processes"(LeDoux and Rogan, 1999, p.269).

In her book *Music as Cognition*, Serafine (1988) comments that "close consideration of all or nearly all of our knowledge about the perception of sound reveals very little that is actually relevant to musical questions"(p.26). She is particularly scathing of scientific approaches that have treated music as sound—that is as acoustic manifestations of vibratory energy—claiming that "in its baldest form ... [such approaches are] rather like searching for the nature and causes of painting and sculpture in the characteristics of pigment and clay"(p.24). Serafine certainly has a point, in that the work of scientists such as Plomp (1976) seems by current standards to be remarkably immune to the influences of human cognition, culture and musical style. On the other hand, in our current intellectual culture of relativism and post-modernism, it is easy to forget that on a very fundamental level, music *is* sound: it may be much more besides, but in a very real sense, it *is* the acoustic manifestation of vibratory energy. Likewise, as the epigraph above suggests, even though the perception of music involves complex high-level processes only some of which are currently understood, it also involves the operation of very low-level processes which detect and respond to acoustic stimuli and to simple patterns. It might be expected that these low-level processes are involved in a fundamental way in the evocation of emotions in response to music.

This chapter considers the perception of sounds and patterns as they appear in a natural acoustic environment and demonstrates how, independently of other higher-level issues, such sounds and patterns can evoke emotional responses by virtue of their acoustic characteristics and the functioning of human perceptual systems. Further, it discusses the role played by sounds, patterns, and the emotional responses that they can evoke in the context of music. The central claim of the chapter will be that our innate, very low level responses to the fundamental characteristics of sounds and basic patterns are integral to any musical listening experience. Whilst such responses alone cannot account for the complexities of emotional response to music, at least a description of them must form a crucial and fundamental part of any listener-centric model of emotional response to music.

Response to Sound

A gunshot in a crowded street provokes an extraordinary reaction: every single person freezes, then tries desperately to assess what has happened. When realisation dawns that the noise was not in fact a gunshot but the sound of an old car back-firing, each person, with a mild sense of relief, goes about his or her business as though nothing has happened. An immediate startle response such as that exhibited at the moment the shot is heard does not rely on every person having carried out a detailed conscious analysis of the spectral quality of the sound and a thorough appraisal of what its significance might be; it is not the result of an indexical association between that sound and fearfulness. Instead, the reaction is a direct response to the sound itself—its intensity, its energy spectrum, its unexpectedness—and a recognition that the sound's source may pose a threat to survival.

A sound is a sensory stimulus, just like a sight, an odour or a touch. In common with all other sensory stimuli, sounds can vary in their intensity level from just perceivable to painfully intense,[1] but unlike a sight or an odour, an intense sound cannot be avoided: we cannot "hear the other way" or "hold our ears".[2] Whatever the precise nature of a stimulus such as a gunshot and regardless of how complex its waveform, the intensity alone is enough to provoke an involuntary physiological reaction.

Researchers who have studied responses to intense auditory stimuli have consistently

[1] In auditory terms, these extremes can be quantified. The average human can perceive a sound whose intensity is 10^{-12} Watts per square metre; a sound of 10 Watts per square metre would be painful to the ears (Pierce, 1983).

[2] Sound is also literally a more "immediate" stimulus than the other modalities: the time taken for neural transduction is noticably quicker for acoustic than for visual stimuli (see Pöppel, 1997).

found that these cause intense physiological arousal and prepare the body for action. An experiment by Jaskowski, Rybarczyk, and Jaroszyk (1994) illustrates the phenomenon clearly: participants were presented with a series of tones that varied only in intensity; the task was to push a button as quickly as possible after hearing each tone. Results showed a negative correlation between reaction time and stimulus intensity; in other words, participants consistently displayed shorter reaction times in response to louder stimuli. In addition, this experiment measured auditory evoked potentials—changes in electrical activity in the brain caused by auditory stimuli—and found that delay between the stimulus and the onset of the potential (EP latency) decreased with increasing stimulus intensity. The following year, a similar experiment by Jaskowski, Rybarczyk, Jaroszyk, and Lemanski (1995) found a positive correlation between stimulus intensity and force in a similar reaction-time task: participants consistently applied more force in response to louder stimuli than to quieter ones; the louder stimuli caused more physiological arousal in the subjects which, presumably, caused them to hit the button harder.[3] These findings fit well with many theoretical models of arousal, which assume that intense stimuli cause instant arousal to prepare the body and mind for action: "intensity increases alertness, which ... leads to faster responses to more intense stimuli"(Nissen, 1977, p.350).[4]

Both of the experiments described above used simple tones as stimuli that varied only in their intensity. In the real world, however, sounds are not simple tones, and our responses to them are not restricted to detection of and reaction to intensity level. Sound may be energy, but that energy can be distributed across the frequency spectrum and over time in an infinite number of ways; that is to say, sounds can have different timbres. The sound of a gun, for example, is not just a loud and unexpected noise; it is a loud and unexpected noise with a distinctive energy spectrum that has a specific and emotionally arousing significance. Unlike intensity, timbre is a multidimensional phenomenon governed by a combination of intensity, frequency and time; as such, it is notoriously complex to describe and understand (Hajda, Kendall, Carterette, and Harshberger, 1997). Despite this, humans are able quickly

[3]In this same study, Jaskowski et al. (1995) noted that "whilst the effect of stimulus intensity is evident for auditory stimulus, we could not find such a relationship for visual stimulus"(p.60). It appears that regardless of the intensity of the visual stimuli, subjects applied a similar amount of force when hitting the button. Such a finding is consistent with the notion that hearing, a sense that tells us what is happening in the world—or at least where energy is being expended and with what force—is somehow more "fundamental" or perhaps "inescapable" than vision.

[4]Nissen (1977) provides a comprehensive, albeit now outdated review of the extensive empirical research in this area.

and reliably to extract much information about a sound source from its timbre and, as will be seen, we are able to do this automatically and non-consciously; just as intense sounds can cause automatic physiological reactions, so can sounds of certain timbres.

In his comprehensive review of timbre research, McAdams (1993) remarked that "humans have a remarkable ability to understand rapidly and efficiently aspects of the current state of the world around them based on the behaviour of sound-producing objects"(p.146); we can differentiate faultlessly between different timbres (e.g. Grey, 1977) even when judgements have to be based on steady-state spectra that lack temporal information (Singh and Hirsh, 1992). Although most studies of timbre have concentrated on perception of musical sounds—specifically those produced by pitched instruments—there exists plenty of evidence indicating that we can easily differentiate complex sounds of a non-musical nature too. For example, Freed (1990), who presented listeners with sounds made by a series of mallets with heads of differing materials, demonstrated that listeners could reliably rate the "hardness" of a mallet from the auditory cues alone. In another study, Li, Logan, and Pastore (1991) found that subjects could even reliably judge (with approximately 70% accuracy) whether a footstep belonged to a man or a woman, basing their judgements solely on auditory cues!

Ability to discriminate timbre could be seen largely as a cognitive issue and of no particular relevance to a discussion of emotion if it were not for the fact that people display a remarkable ability not only to differentiate between timbres but also to assess the significance of a sound in emotional terms and in terms of its likely significance to their own well-being.[5] Further, they appear to be able to do this extremely quickly. For example, Nielzen and Olsson (1993) synthesised a set of complex sounds based on natural spectra, each of which lasted for approximately three seconds. They found that experiment participants were consistently able to rate the sounds according to dimensions of tension/relaxation, lightheartedness/gloom, and attraction/repulsion; perhaps a more impressive finding was that when each stimulus was cut so that it sounded for a mere half a second, ratings were just as robust.[6] Unfortunately, this experiment is not without its problems: due to its reliance on

[5]A timbre produced by a recognised object might evoke emotion by virtue of association if the obect producing it has emotional significance. This is emphatically not the issue with which we are currently concerned, however. Here, the phenomena of interest are acontextual responses to acoustic stimuli in and of themselves. Emotional significance imbued by virtue of association will be discussed in Chapter Four.

[6]There were, in fact, small differences in some of the ratings but the overall pattern of results remained consistent between 3 second and 0.5 second stimuli.

introspective judgements, it may tell us more about participants' belief structures than about their direct responses to the sounds.

Luckily, despite this shortcoming, the findings are corroborated by a wealth of research that has investigated the perceived urgency of different sounds using non-introspective methods; this research suggests differentiated emotional response to sound at a pre-conscious level. For example, a particularly interesting investigation has been carried out by Burt, Bartolome, Burdette, and Comstock (1995) that attempted to assess the extent to which participants could attribute urgency to auditory warning signals. The purpose of the investigation for them was to develop sophisticated warning systems for pilots that would help catalyse action in an emergency situation whilst not provoking unnecessary alarm in rather less dire circumstances: "by providing auditory warnings that are designed with sound parameters capable of matching a pilot's psychoacoustic perception of urgency, warning urgency level can be quickly discerned, and time and resources may be saved during critical phases of flight"(p.2329). Such warning systems would not rely on a pilot's ability to recognise a sound and to remember what it represents in some sort of "sonic catalogue", but could instead exploit innate, automatic human responses. Presumably, if a pilot actually perceived a sound to be innately "urgent", his or her reaction time in response to it would be greatly reduced. The experiment that Burt et al. devised involved measuring the degree to which various auditory warnings affected participants' reaction time and brain activity (auditory evoked potentials) whilst performing a tracking task; the degree to which reaction time was affected by each sound was assumed to be indicative of the degree of perceived urgency.[7] Results showed that level of perceived urgency of a sound could be manipulated by altering several acoustic parameters, most notably the fundamental frequency and energy spectrum, whose effect was independent of intensity. In other words, some noises sound innately more "urgent" than others.

Equally interesting was an additional finding that participants could be trained to override their initial responses, and learn to associate specific sounds with a certain level of urgency, thereby changing their default response patterns. It would appear, then, that response to the timbre of a sound may be qualitatively different from response to intensity level: although no research has examined this explicitly, it would seem incredibly unlikely

[7] This assumption seems justified in the light of research such as that of Jaskowski et al. (1994, 1995).

that someone could be trained to learn a counter-instinctive response to the *intensity* of a sound stimulus; these responses are just too ingrained. The work of Burt et al. implies that timbral quality of a sound can acquire an emotional significance by at least two routes: the spectrum may be perceived to be innately emotional—in the same way that an intense sound is more arousing—or a particular sound may acquire a significance through learning.

If such learning is important for well-being, it may not be "learning" in the usual, high-level cognitive sense of the word; it may a phenomenon quite distinct from the indexical association mentioned in the previous chapter (and discussed more fully in Chapter Four). Human clinical reports and recent research on animals has indicated that we possess a special kind of memory, an "emotional memory " (LeDoux, 1998) that works independently of declarative memory. For researchers such as LeDoux, such phenomena can only be understood by examining the human or animal brain. Claperede (1951, cited in LeDoux) reported a fascinating case of a brain-damaged patient who, despite being able to remember the past, was unable to store new memories. As a result of her brain damage, whenever the patient went for a consultation, she acted as though she had never seen her doctor before: she would always shake his hand and introduce herself as if at their first meeting. After several consultations, Claparede tried what would now be called a "conditioning" experiment: he attached a pin to the palm of his hand so that the next time the patient came to introduce herself, she pricked herself on the pin and recoiled with pain. On future visits, although still insistent that she had never met her doctor before, the patient refused to shake his hand; remarkably, she was totally unable to explain why she acted in this way. "Claperede was seeing the operation of two different memory systems in his patient—one involved in forming memories of experiences and making those memories available for conscious recollection at some later time, and another operating outside of consciousness and controlling behavior without explicit awareness of the past learning"(LeDoux, 1998, p.181).

An experiment on rabbits by Jarrell, Gentile, Romanski, McCabe, and Schneiderman (1987) tells us more about the nature of such fear-conditioning, particularly in response to auditory stimuli. In this experiment, the rabbits were conditioned to expect that one type of sound would be paired with an electric shock whereas another would not. After a while, these rabbits would consistently display fear when the former sound was presented, but showed no reaction when presented with the latter one; whereas one sound had become a highly-charged stimulus, the other had not. Later, the rabbits were given cortical lesions, so that they effectively lacked higher-order cognitive capacity. When presented with the stimuli again, they could not differentiate between the two sounds but responded fearfully to *both* of them. Having lost their ability to discriminate between the timbres, they followed what might be considered to be the safest evolutionary strategy: they prepared themselves physiologically for the worst. Whilst differentiation of timbres may require a cortex, the ability to associate a sound with a basic emotional reaction, such as fear, is not; it relies on evolutionarily early brain systems.[8]

Recent neuroscientific research has discovered a lot about brain systems involved in the emotional processing of stimuli. It appears that emotional processing of sound mainly involves the auditory thalamus, amygdala and cortex (see e.g. Morris, Scott, and Dolan, 1999)[9] with fear processing centred mostly around the amygdala (e.g. Scott, Young, Calder, Hellawell, Aggleton, and Johnson, 1997); the amygdala is known to be involved in mediating brain-stem activity, which governs physiological arousal. Cortex lesions do not affect fear conditioning at all, but thalamus lesions do (LeDoux, Sakaguchi, Iwata, and Reis, 1984); the auditory thalamus, one of the first parts of the brain to process sound stimuli, is connected directly to the amygdala as well as to the auditory cortex (LeDoux, Farb, and Ruggiero, 1990). This map of brain connectivity can give us insight as to why we might respond to sounds the way we do. As LeDoux (1998, p.162) wryly notes:

> "Neurons in the area of the thalamus that projects to the primary auditory cortex are narrowly tuned—they are very particular about what they will respond to. But cells in the thalamic areas that project to the amygdala are less picky—they respond to a much wider range of stimuli and are said to be broadly tuned. The Beatles and the Rolling Stones ... will sound the same to the amygdala by way of the thalamic projections but quite different by way of the cortical projections".

So an auditory stimulus can cause brain activity in the amygdala without any cognitive processing occurring at all, but cognitive processing is required to make fine-grained distinc-

[8]The extent to which such findings can be phylogenetically extrapolated to the human brain has been questioned by researchers such as Grossenbacher (1996), who notes that older brain systems could have adapted to perform quite different functions in humans. However, researchers such as LeDoux use a panolpy of neurophysiological and clinical observations to defend the stance vehemently.

[9]But see also Adolphs and Tranel (1999) who claims that the amygdala is not at all involved in the detection of emotion in prosody.

tions and mediate the amygdala response. An intense auditory stimulus, or a sound with certain general timbral characteristics will cause an amygdala response directly, to prepare the body for action; higher-level cortical processes that have access to rather more information about the nature of the stimulus and its likely significance can effectively veto the action, informing the amygdala that the stimulus is unworthy of attention! As LeDoux (1998, p.165) points out, whilst this two-tier system may seem rather odd, it is a very efficient mechanism for ensuring quick responses to life-threatening situations:

> "Although the thalamic system cannot make fine distinctions, it has an important advantage over the cortical input pathway to the amygdala. That advantage is time".

We are now in a position to return to the scenario set out at the start of this section, namely the gunshot in a crowded street. Everyone in the street hears a loud bang; it is an intense sound with a broadly recognisable spectrum. Immediately, auditory thalamus projections to the amygdala fire, and the body prepares itself for emergency action. Meanwhile, the stimulus reaches the auditory cortex, which recognises that the spectrum is not quite right for a gun; this realisation, coupled with the sighting of an old car spluttering down the road, allows the sound source to be identified as the back-firing car. Cortical projections to the amygdala send this new analysis on, and the body's physiological activity is allowed to subside. Had the immediate amygdalic response not occurred and had the sound belonged to something rather more ominous, the time taken for cognitive analysis might have rendered any chance of survival impossible.

Discussion so far has largely focused on intense or fear-inducing stimuli, with a brief discussion of brain systems that might be responsible for our responses to them. There are two reasons for this: first, these cases constitute good, clear examples of auditory stimuli which provoke emotional reactions independently of context and higher-level concerns;[10] second, whilst there is a plethora of research on auditorily-evoked fear and response to intense signals, there is a comparative dearth of literature on emotional response to non-intense sounds outside of a musical context. Not all sounds are intense or fear-inducing, however, and it

[10]Physiological arousal is not, of course, synonymous with emotional response, as Schachter and Singer (1962) famously illustrated. As they also demonstrated, however, very little is needed to turn physiological arousal into a fully-fledged emotional response, and that which is needed can be trivially provided by cognitive context. For the purposes of the present discussion, therefore, evidencing that an acoustic stimulus causes physiological arousal is tantamount to evidencing the evocation of emotional response.

seems reasonable to suggest that non-intense sounds can also evoke an emotional response. According to Frijda (1986, p.264) "stimuli of weak of moderate intensity induce pleasant emotions, and intense stimuli induce unpleasant emotions", which suggests that an examination of moderate intensity sounds might be instructive.

Unfortunately, only one study to date appears to have concentrated on emotional responses to sounds that are neither necessarily of high intensity nor of particularly arousing timbre, namely an investigation by Sakamoto, Hayashi, Tsujikawa, and Sugiuri (1997) on physiological responses to average sounds from our daily lives. In this study, participants were presented with a series of sounds that occur frequently in daily life, such as the sound of running water, voices and car engines; the task was to identify and rate each stimulus whilst the experimenters measured pulse and blood pressure level. Results showed that people displayed physiological reactions to positively valenced sounds, but the study said little about the nature of the relationship between sound and response. We do not know, for example, whether the sounds evoked reactions directly or whether the observed physiological arousal was the product of sophisticated cogitation. There is clearly scope for further research in this area.

One class of sound that has received a lot of attention in the psychological literature, but has not yet been mentioned here, is that produced by the human voice: an abundance of evidence suggests that we are extremely adept at recognising and responding to complex emotional signals encoded within the voice, and that this aptitude, which is not reliant solely on spectral content and intensity but also on prosody and larger-scale patterns, may be as innate and pre-conscious as our responses to intense signals, or those of a certain timbre. Evidence suggests that human vocalisations are perceived and responded to quite differently from other classes of sound, and that the relationship between such sounds and music is also very different. In fact, the difference is such that emotional response to human utterance is treated in this thesis as a fundamentally distinct phenomenon, and will be addressed in Chapter Three.

In summary, we can respond to sound—to auditory stimuli—on a very low level. Intense sounds, those with certain spectral characteristics, or those that have become associated with danger, cause the activation of phylogenetically early brain systems, which in turn necessarily causes physiological arousal. This occurs in spite of, and in addition to any higher-level responses to those sounds.

Response to Pattern

Discussion so far has considered individual sounds as isolated entities. Very few noises in the real world consist just of single impulses of a specific spectrum, however. Instead, sounds come in groups—a *series* of footsteps, *gusts* of wind—and those groups are part of a complex auditory environment. McAdams (1993) may have commented on our ability to discover much about the world from the sounds which emanate from it; much of that discovery, however, comes not so much from recognition of isolated acoustic stimuli, but from our ability to disentangle the patterns of a complex world. We may well have an innate ability to detect and respond to the significance of single sounds, but this often relies on an equally innate ability to detect and respond to patterns in the auditory environment.

The first psychologists to study pattern perception in a systematic way were the so-called Gestaltists, a group of German researchers who emigrated to New York between the two world wars (Eysenck and Keane, 1995). They noted an apparent human pre-disposition toward grouping objects in a visual scene according to a simple set of laws: objects are grouped if they are close together in space, if they are similar in appearance, if they display good continuation (i.e. an imaginary line drawn between them has few interruptions or changes of direction), or if they imply closure (i.e. form an incomplete part of, say, a circle); in addition, objects that move in synchrony form a perceptual grouping.[11] Although the theory that the Gestalt psychologists provided by way of explanation for their observations has been shown to be somewhat flawed, the phenomena that they describe appear consistently and universally. Therefore, they "have not been seriously challenged or undermined over the years"(Eysenck and Keane, 1995, p.34).

Whilst the Gestaltists were primarily interested in visual perception, the observations that they describe have analogies in the auditory domain. For example, sounds which are close to one another in frequency might be said to be grouped by "proximity"; sounds which have a similar spectrum might be said to be grouped by "similarity"; a set of sounds each of whose fundamental frequencies is higher than the preceding one might be said to be grouped by "good continuation"; two or more sounds that change in frequency together

[11] An authoritative explanation of Gestalt psychology can be found in Koffka (1935).

could be considered to be sharing a "common fate". This similarity between the two perceptual domains is something that the Gestalt psychologists themselves noted (see e.g. Koffka, 1935), and has been considered frequently by audition psychologists since (see e.g. Darwin and Carlyon, 1995). In recent years, various phenomena relating to auditory pattern perception—including those that are analogous to the Gestalt laws—have been studied extensively. Most notably, Bregman (1990) combines a comprehensive review of research in the field with a theory of significant explanatory power.

Bregman realised that examples of auditory perception that can be described by Gestalt principles are not isolated phenomena but are part of a larger scheme. It may be true that we can hear a series of individual sounds and bind them into perceptual groups, but the reality in a world full of sound sources is that the difficulty is not so much in the grouping of isolated entities but in segregating the complexities of the auditory environment into distinct perceptual streams. The problem of "auditory scene analysis", to use Bregman's term, is to analyse the constant stream of acoustic information arriving at the ears, in order to reveal information about the events that are taking place in the environment:

> "Sound is created when things of various types happen. The wind blows, an animal scurries through a clearing, the fire burns, a person calls. Acoustic information, therefore, tells us about physical 'happenings'. Many happenings go on at the same time in the world, each one a distinct event. If we are to react to them as distinct, there has to be a level of mental description in which there are separate representations of the individual ones"(Bregman, 1990, p.10).

Sounds arising due to an animal scurrying through a clearing or fire burning are not single, isolated impulses but patterns that are perceived as a single stream; one could add to the list a series of footsteps, or the twitterings of a bird. To take a musical example, the hearing of a series of notes played by a single instrument as a melody could be attributed to the gestalt principles of similarity (all the notes have similar timbre), promixity (assuming the notes all fall within a reasonably narrow pitch range) and good continuation (provided that the melody is not too angular); a scene-analytical explanation would add that a listener has used cues such as similarity, proximity and good continuation to segregate out a sound source (the flute) from the complex auditory environment. Although a detailed discussion of Bregman's theory and the nature of auditory scene analysis is well beyond the scope of this thesis,[12] the idea that we can segregate the auditory environment into streams (or conversely, that we make perceptual groupings out of parts of the acoustic environment) and that the resulting streams (or groups) constitute perceptual units that can be considered or recognised, is crucial to present concerns. Although identification of a sound source might be dependent on learning—it is impossible to identify a series of crunches as the sound of a sabre-toothed tiger without having encountered one before—the ability to segregate (or group) sounds appears to be innate and automatic (e.g. Bregman, 1990; McAdams and Bertoncini, 1997). Different environments may sport their own peculiar sounds, but the general physical properties of sounds remain constant, and as Bregman (1990, p.39) observes: "the internal organs of animals evolve to fit the requirements of certain constant factors in their environments. Why should their auditory systems not do likewise?".

As discussed in the previous section of this chapter, we respond emotionally to (or at least are aroused by) loud sonic impulses and acoustic signals with certain spectra. It is not difficult to see, in addition, how the process of auditory scene analysis might lead to similar arousal. If we have an innate ability to segregate sounds into auditory streams, and that ability derives from analysis of certain constant physical properties of sound, then hearing patterns in the auditory environment will involve the generation of expectancies as to how those sounds will continue or change over time. For example, a series of footsteps getting quieter and quieter (implying that some animate object is moving further and further away) would generate the expectancy of more, even quieter footsteps at a similar rate that eventually would subside completely; if, suddenly, the footsteps doubled in speed and became dramatically louder, then the expectancy would be violated. It might be expected that such a violation would cause an immediate physiological reaction: not only would such a reaction be predicted by arousal theories of emotion (e.g. Mandler, 1984), but it also fits well with LeDoux's work on fear in hypothecating an emotion system that reacts quickly to anything untoward, only moderating its response later (if necessary) once enough time has passed for a full cognitive assessment; a mechanism supporting such a reaction would be simply a system for survival.

One of the remarkable features of human perception is its ability to avoid uncertainty by filling in missing information. Bregman (1981) provides a visual example with a series

[12] Any reader who wishes to explore this further is encouraged to read Bregman (1990).

of strange-shaped fragments: when seen in isolation, the fragments are meaningless, but when another masking shape is overlaid, it is suddenly apparent that the "fragments" in fact form collaged images of the letter "B". As soon as the strange shaped contours are perceived as being attached to the mask rather than the fragments, the perceptual system can "fill in" what has now simply become missing information.[13] An auditory example of an analogous phenomenon is provided by Dannenbring (1976): a siren-like sound sweeping up and down the frequency range is interrupted at regular intervals with a noisy mask; even though listeners cannot hear the siren behind the mask, they describe the siren as sounding continuously. If the mask is replaced with moments of silence, they hear the siren as being constantly interrupted—as there is no hidden information, the illusion of continuity cannot be maintained. In short, our perceptual systems seem to be optimised to avoid ambiguity and uncertainty whenever possible.

There are always times, however, when uncertainty cannot be avoided. There may exist strange series of shapes that cannot be explained by masking phenomena, or series of patterns that appear not to obey standard physical laws; in fact, for any number of reasons, there may be aspects of the auditory environment whose sources we cannot ascertain. The resulting uncertainty in such a situation can constitute a source of physiological and emotional reaction. In his discussion of human anxiety, Lazarus (1991, p.234) observed that "to survive, humans need to impose meaning on events in a confusing world"; this observation seems valid at many levels of perception and cognition. When events cannot be interpreted in the face of a possible threat to survival, emotions are triggered. As auditory perception provides information about the immediate auditory environment and is processed, for the most part, outside of conscious intervention, *any* uncertainty could pose a threat to survival; therefore, inability unambiguously to interpret an auditory pattern should necessarily cause physiological arousal which, in turn, can be experienced as emotional response.

An obvious route by which the processing of auditory patterns might lead to an emotional reaction is if those patterns are consciously recognised as belonging to emotionally-valent stimuli. For example, if a series of footsteps is recognised as belonging to a sabre-toothed tiger, or a series of shrieks is identified as a man-eating cormorant, this in itself should be enough to evoke a strong reaction, simply by virtue of the stimulus. Consid-

[13] This illusion can be seen as the operation of gestalt principles or of auditory scene analysis at work; for details, see Bregman (1990).

ering Bregman's argument that auditory scene analysis has a function, however, and that the function is to tell us about happenings in the world, another route becomes apparent: the nature of an auditory pattern in itself could imply a sound source significant enough to cause alarm even if that pattern has never been heard before; such a stimulus may be perceived to be emotionally valent in and of itself. For example, a series of crashes, steadily increasing in amplitude, might of itself imply a large animate object moving rapidly closer, posing an ever increasing risk to the listener's safety. Within virtually any extant theory of emotions—certainly any of those mentioned in this thesis—such a situation would cause arousal followed, hopefully, by evasive action.

Footsteps or the crashes of a sabre-toothed tiger do not always get closer and louder; sometimes, they or equivalent sounds can form a steady pattern of moderate intensity that is neither particularly arousing nor alarming. Such sounds belong to a special class of auditory patterns, namely, rhythms. Quite independently from moment-by-moment appraisal of the auditory environment, survival, and the fear response, there appears to be something about these regular, repeating patterns that empowers them to evoke physiological effects that are qualitatively different from the immediate, survival-instinct reactions that have been the subject of most of this chapter; whereas the sounds and patterns discussed earlier lead to immediate physiological arousal and action-readiness, the main effect of rhythmic patterns is to mediate mood (see e.g. Rigg, 1940, as cited earlier), affective phenomena that are longer-lasting and more stable than short-lived physiological reactions and concomitant emotions (Davidson, 1994).[14] It has often been suggested that the effect of regular rhythmic patterns on mood is related to the importance of regular rhythmicity in physiology (see e.g. Fontana and Loschi, 1979). Heart-beats, breathing rates and circadian rhythms are all internal regularities by which our physiological lives are controlled; externally produced rhythms can affect the speed of these physiological phenomena, thus altering base arousal level, and thereby our general level of awareness (Wallin, 1991). The powerful effect of rhythmic auditory stimuli is clearly illustrated by Smith and Steinschneider (1975), who carried out a study of the sleeping habits of newborns and the effect of the noise of mothers' heartbeats, finding that not only did a simulated heart-beat sound help newborns sleep but also that "babies born to low-heart-rate mothers generally fell asleep much faster, slept more, and cried less than

[14]See Chapter Four for a rather fuller discussion of the difference between emotion and mood.

babies born to high-heart-rate mothers"(p.577).

The difference between our response to isolated sounds and patterns and to steady rhythm is explainable within the context of auditory scene analysis: a pattern of regular repetition, once initiated, contains no new information about the environment; provided it does not suddenly change (causing an expectancy violation), and provided its intensity or timbre is not so extreme that every pulse is felt individually as an intense stimulus, it contains no emotionally charged material.[15] Instead, a regular pattern is an indicator of the activity level of an environmental sound source; it tells us *how much* is going on and, perhaps as a corrollory, how generally aware—or aroused—we need to be. As such, a rhythm provides a steady background against which other happenings can occur.

In the light of these observations, it is perhaps unsurprising that rhymic patterns are used extensively in therapy (see e.g. Keats, 1995) and by individuals and groups for mood induction (see e.g. Hakanen, 1995); sometimes, their effect on physiology has even been employed to evoke hypnotic trance (see e.g. Akstein, 1973). It is equally unsurprising that many cultures use rhythmic patterns as part of rituals (see e.g. Rouget, 1980). As Clynes and Walker (1982, p.174) put it, "in the physical world, rhythm is fundamental to existence".

In summary, then, individual sounds can evoke emotional responses by virtue of their status as intense or significant stimuli. Often, sounds do not appear in isolation, but rather exist in groups and patterns as part of a complex auditory scene; patterns can evoke emotions either by virtue of their significance, or by confusing or surprising our perceptual systems. Emotions evoked by isolated sounds and auditory patterns are strongly linked to physiological arousal, can be sparked without conscious awareness, and are fast-acting. By contrast, regular rhythmic patterns that do not consist of intense or significant stimuli are able to mediate mood, possibly by virtue of their similarity to the internal physiological rhythms by which our lives are governed.

Sounds and Patterns in Music

So far in this chapter, we have seen that sounds can cause physiological arousal by virtue of their energy spectrum and intensity, that patterns can evoke emotional responses due to

[15] Even if a pattern consists of individual impulses which are of extreme intensity and timbre, physiological arousal tends to subside after the initial shock, because we become habituated to intense acoustic stimuli (e.g. Putnam and Roth, 1990).

their internal structure and its relationship to human perceptual systems, and that rhythms mediate mood and baseline physiology. There is no reason to believe that these effects will be any less apparent if the sounds are heard within a musical context; in fact, at the level being discussed here, empirical evidence and common sense converge in suggesting that people will respond to the sounds of music in exactly the same way as they do to all the other sounds in the world. A loud, sudden bang is still a loud sudden bang whether its source is an explosion in a crowded street or a bass drum in a symphony orchestra. A vital feature of the physiological response to intense stimuli is the fact that it is fast-acting and not particularly discriminate; only once the body is ready for action is there time to work out the details of what actually *caused* an intense sound and to retrospectively realise it may not have been cause for alarm; if this description of response characteristics is valid—and overwhelming evidence suggests that it is—it must *by definition* apply to intense sounds in music just as much as anywhere else. Similarly, a repetitive rhythm is a repetitive rhythm whether it belongs to the footsteps of a sabre-toothed tiger, the beating of a human heart, or the drum track of a song.[16] In short, the sounds and patterns present in music can cause arousal, evoke emotion and mediate mood on a level at which the musical context of the stimuli is utterly irrelevant.

Of course, there are differences between our responses to sounds in the general auditory environment and our response to those same sounds in music: people do not run screaming out of a concert hall when the canons are fired in Tchaikovsky's *1812 Overture* just as they do not flee from the cinema when shown footage of a huge green monster approaching the front of the screen; by contrast, they would probably run extremely fast if encountering either the sounds of canon fire or a huge green monster in real life.[17] The reason for this distinction is that the significance of a loud sound of certain timbre, or a rhythmic pattern of a certain tempo is quite different in a musical and a non-musical context: people realise that the sounds of music or screened images of monsters are coming from a "safe" source. In a musical context, the canon fire takes on its own meaning as the creation of a musician (or composer in the western art music tradition). In this light, sounds—even emotive ones—

[16]This observation is vindicated in part by a study by Panksepp (1997) who attempted to measure differences in cortical arousal in response to "happy" and "sad" music. He found that the modest differences in EEG activity that could be detected were equally evident in response to non-musical auditory stimuli.

[17]The huge green monster example is adapted from Walton (1990).

are just a part of a deliberate musical act and heard as such.[18] Such differences, however, manifest themselves at much higher cognitive levels than the types of responses that are the subject of this chapter; at the level of auditory thalamus projections to the amygdala, the canons of the *1812 Overture* are merely extremely loud bangs, are heard as such, and are responded to as such. The remainder of this chapter is devoted to outlining the role played by low-level responses to sounds and patterns in the musical evocation of emotion. Evidence will be presented not only to suggest that our emotional response to sounds and patterns in music can, at a low level, be understood as an identical phenomenon to our response to sounds in the acoustic environment, but also to illustrate that emotive properties of sounds are deliberately exploited by musicians and, in western art music, have become intrinsic to composition. In the process, it will hopefully become apparent that much empirical research on response to intensity, timbre and pattern in music can be seen not as observations of isolated phenomena, but rather as illustrations of a more general response phenomenon, namely our sensitivity to the sounds and patterns of our environment.

To include loud noises in a composition, such as the canon in the *1812*, is literally to "borrow" the emotion evoking properties of the noises themselves and to use them as is; the effect is hardly a subtle one, and would work equally within or without a musical context. In fact, it works specifically because of the irrelevance of musical context. Often, however, low-level arousal effects of sound are integrated into a musical structure in far more subtle ways, a particularly fine example being the way in which composers in the western art music tradition control intensity. The rate of change in the intensity of any acoustic stimulus provides important information about the sound source: specifically, a stimulus that increases in volume is getting closer or more energetic; it is almost certainly one to which it would be wise to pay attention! The ramifications in a musical context are that an increase in dynamic level should lead to an increase in listener arousal and attention; a natural conclusion to draw from this is that a composer wishing to arouse an audience, or just to keep its attention, should perpetually increase intensity as a work unfolds. Research by Mathews (1979) corroborates this suggestion by demonstrating that the perceived strength of a musical crescendo is related to its length.

[18] The importance of context in mediating emotional response to music will be discussed further in Chapter Four. For a lengthly exposition on the cognitive significance of perceiving something to be a deliberate communicative act, see Sperber and Wilson (1995), particularly their consideration of "acts of ostention".

Unfortunately, perpetually increasing intensity throughout a work is not necessarily a practical proposition for a composer writing a long piece of music, especially if resources are limited; yet there exist countless pieces of music that seem to maintain their strength using exactly this technique. This apparent paradox seems to have captivated Huron (1990a,b, 1991, 1992) who suggested that in these situations, it is in fact an alternative but equivalent strategy that is adopted:

> "Where a musical genre makes available many resources, it is possible to engineer long passages where the amount of stimulation is constantly increasing. However, where the music genre provides only modest resources, the possibilities for a perpetual escalation of stimulation are soon exhausted. In such circumstances it is possible to maintain attention by rapid and large changes of dynamics, or by structuring the work as a sequence of stimulus 'ramps'"(Huron, 1992, p.89).

These *ramps* involve a gradual building of dynamic level followed by comparatively quick subsidence, so that the intensity is increasing by an optimal amount for the majority of time. Remarkably, extensive analysis of several works found such a pattern of gradual crescendo and quick diminuendo to be ubiquitous in the western art music repertoire, particularly in musical genres with limited "possibilities for a perpetual escalation". In effect, the power of an increasing-intensity sound stimulus has been adopted and turned into a musical technology by a vast number of composers. In the abstract, it would be difficult to interpret Huron's findings as anything other than consistent use of a clever compositional device within a repertoire; here, however, they can be seen as an example of composed music deliberately exploiting emotion-evoking properties of auditory stimuli.

A similar argument can be made for the timbral characteristics of sounds; composers choose instrumentation for their music according to the effect they wish to achieve (see e.g. Piston, 1955) and musicians subtly adapt the timbre of their instruments by modifying articulation and through techniques such as vibrato (see e.g. Gabrielsson, 1999, for a review). Although the range of possibilities is limited in both cases—composers tend only to score for extant instruments whose sounds have a loosely defined spectrum and instrumentalists can usually alter the spectral characteristics within a limited range—the result is music that draws deliberately on emotional responses to specific classes of sounds. Recent technological advances (i.e. fast, cheap computers) have allowed musicians far more control over the spectrum of sounds they produce than they had in the past, and have led to an increasing interest in timbral dimensions of music. Electro-acoustic composers such as David Wessel have designed software that allows exploration of a continuous "timbre-space", and although such work tends to be somewhat arbitrary in its choice of dimensions for "timbre-space", it is not hard to imagine composers consistently using timbral techniques whose effects would be equivalent to Huron's stimulus ramps as psychologists discover more about our perception of timbral spectra. A more prosaic but nonetheless fascinating example of timbral manipulations can be seen in the popular music recording studio, where it is commonplace to employ combinations of reverberation, equalisation and compression to make the sound of a song more dramatic, more excited or more urgent. Whilst such manipulation of recorded sound remains a black art with no absolute rules (Ross, 1987), it is abundantly clear that record producers and engineers deliberately harness the emotion-evoking properties of certain timbres when creating their recordings.

As has already been noted, a relatively large amount of research has investigated the perception of timbre, most of which has been concerned with musical timbres. This research has both revealed that we are rather good at differentiating between timbres, and suggested that timbres can in themselves evoke emotional responses. Unfortunately, comparatively little research has investigated emotional responses to musical timbre in a systematic way, possibly owing to the general difficulty of quantifying something that "is an emergent property that is partly a function of the acoustical properties and partly a function of the perceptual processes" (Handel, 1995, p.458). Hopefully, this is a situation that will change as researchers begin to understand more about the relationship between the acoustic and perceptual aspects of timbre. Meanwhile, we at least have strong evidence that some timbres are perceived as more pleasant than others (see e.g. Halpern, Blake, and Hillenbrand, 1986) and can assume, just as we do for intensity, that those low-level mechanisms that make us respond emotionally to timbre will be operative when we hear music, just as when we hear other, non-musical sounds; certainly people involved in various aspects of the music industry appear to make this assumption.[19]

[19]There is a growing body of research concerned with understanding communication of affective gesture in music, a topic that includes examination of production and perception of timbre (see e.g. Gabrielsson and Juslin,

Supporters of the intrinsic-extrinsic dichotomy and exponents of a listener-centric model of emotional response to music would disagree violently on the importance of the phenomena discussed so far. An exponent of the instrinsic-extrinsic dichotomy might argue that all are merely examples of extrinsic sources of emotion and therefore at most of peripheral interest: if the acoustic properties of a sound—its intensity and its timbre—cause physiological arousal, this has nothing whatsoever to do with music except in a circumstantial sense. Exponents of a listener-centric model, by contrast, would argue that a sound whose acoustic properties necessarily and unavoidably provoke physiological arousal owing to their effect on our perceptual systems must be absolutely central to any listening experience; their ontological status with respect to the music is utterly irrelevant. In support of this claim, they might invoke much of the research cited so far as evidence that musicians themselves are implicitly in agreement, whether or not they know it. By contrast, no such clear water would be found between stances if we consider patterns. Even according to musicologically rooted theory, the way in which we respond to patterns—the workings of gestalt principles and the mechanisms of auditory scene analysis—are intrinsic to the very core of music itself.

Patterns appear in all musical dimensions, such as pitch, time, harmony, and even large-scale musical structures. Just as for the acoustic properties of sounds, there seems no reason to believe that the mechanisms by which patterns in the environment can evoke emotional reactions are not engaged equally when we listen to the patterns of music. In fact, given that any musical style adheres, however loosely, to a set of stylistic norms at many levels, it seems reasonable to assume that listening to music involves the constant generation of expectancies whose confirmation or denial can lead to an emotional response in the listener. It is commonly accepted that the generation, confirmation and denial of expectancies is crucial to the dynamics of Western music; this notion forms the basis of Meyer's (1956) theory and much work on emotional response to music that has been carried out since. Expectations are generated on many levels in a musical structure, and perception of those expectations can be crucially determined by a listener's knowledge of the musical style in question (see e.g. Lerdahl and Jackendoff, 1983): after all, a listener can only generate style-specific expectations if he or she is familiar with the style.[20] We should expect, however, that there exist

1996). This will be considered in Chapter Three.

[20]The significance of musical expectations that involve style-specific knowledge, and the way in which they can evoke emotional responses is considered in Chapters Four and Five.

potential expectations that are independent of style-specific knowledge and, as such, are are generated for *any* listener, regardless of his or her familiarity with a musical genre. Such expectations are those that derive from the automatic workings of human perceptual systems; it is they that are of relevance to present concerns.

As we have seen, humans parse events in the auditory environment according to the Gestalt laws: certain patterns are innately implicative of certain groupings or continuations. Responding emotionally to patterns that disobey these laws is as innate a behaviour as emotional response to an intense or disturbing sound in the acoustic environment. Our automatic response to patterns is highly robust, because expectations generated according to Gestalt principles or the mechanisms of scene analysis work independently of knowledge-based influences. This observation has clear ramifications for music listening, a process that involves both the low-level perception of sound and the interpretation of style-specific structures. According to Narmour (1990, 1992), a dichotomy between "top-down" expectations generated from schematic knowledge of a music genre and "bottom-up" expectations generated by the workings of the human perceptual system is crucial to our perception of music. Noting that "the demands, goals, and functions of bottom-up processing frequently diverge from, and are always independent of, those of top-down processing"(1990, p.53), Narmour proposed that melodies always present both top-down and bottom-up implications, the former based on schematic knowledge of *style structures*, and the latter based on basic *style shapes*, which are pan-stylistic and perceived innately. The two types of implications interact in perception, but crucially, whereas *style shapes* can help form *style structures*, the converse is not true: "the style shapes selected by the Gestalt principles hypothesize the kinds of parametric implications resistant to learning"(p.64). The nature of these implications was the subject of an intricate predictive model.[21] Narmour's suggestion that top-down and bottom-up perception takes place in parallel even lends musicological credence to the idea that these low-level principles that govern perception of the acoustic environment in general are still operable in a musical environment.

Schellenberg (1996, 1997) executed a number of experiments that were designed to test Narmour's highly intricate model. Specifically, he was interested in testing the complex set of rules Narmour invented to predict expectancies that should be set up by *style shapes*.

[21] For a comprehensive review of Narmour's theory, see Cross (1995).

Although results indicated that the model was somewhat over-specified, as a significantly simplified version could be employed with no loss of predictive power, its basic tenets were strongly supported. Of particular interest to present concerns were the findings that listeners responded to implications encoded within style shapes in a consistent and predictable way regardless of their cultural background or the degree to which they were musically trained. Narmour assumed that response to such structures should be impervious to all external influences owing to its reliance on hard-wired features of the human perceptual system; Schellenberg's data vindicates the assumption.[22]

Operation of Gestalt laws and scene analytical principles in music is not confined to the domain of melody. Temporal patterns are equally affected by them (e.g. Gabrielsson, 1993); a rhythmic pattern can generate expectations which can be confirmed or violated as music unfolds on a number of hierarchical levels.[23] In terms of causing arousal or evoking emotional responses, research already cited as to the special physiologically-moderating status of regularly repeating patterns might lead us to expect that rhythm should have a significance in music beyond its role as a source of expectancy generation, however. As a large number of empirical studies illustrate, this assumption holds. Rhythm—or more specifically rhythmic density—is one of the main parameters by which listeners judge level of activity in music; given the relationship between activity and arousal, the potential significance for emotional response should by now be clear. Of course, rhythmic patterns in music have a role that is far more important than simply defining a rhythmic density, but evidence suggests that level of activity—rhythmic density—is the single aspect of rhythm that listeners notice the most. Gabrielsson (1973) carried out an extensive study of emotional responses to rhythmic patterns using both musical and non-musical stimuli. Participants in the study were asked to listen to the patterns and rate them according to approximately one hundred preselected adjectives.[24] Factor analysis of the results revealed that perception of the rhythms could be modelled using a mere three dimensions: vitality or dullness, excitedness or calm, and rigidity or flexibility; of these, the most important contributors were vitality and excitedness. By

[22] A study by Schmuckler (1989), which predates Narmour's model, found that melody and harmony worked independently in the generation of listener expectations. This finding further supports the claim that response to style shapes remains impervious to the influences of even rich and complex potential contaminants.

[23] Listeners' sensitivity to expressive timing deviation in performance (see e.g. Repp, 1997) stands testament to the extent to which expectancies can be manipulated in the temporal domain.

[24] These adjectives were chosen from a much larger list of words that were compiled with the help of a number of musicians, whose remit was to suggest as many words as possible that could be used to describe rhythmic phenomena.

far the most significant factor in determining these ratings was the inter-onset interval of the individual pulses; in other words, it was rhythmic density. This finding does not stand alone: Gundlach (1932), for example, noted that listeners tend to perceive fast tempos—high rhythmic density—as animated and slow tempos—low rhythmic density—as dignified or calm.

Whilst the above studies provide evidence that listeners notice rhythmic density in music and ascribe level of activity to it, they tell us nothing about emotion directly. More recently, however, this gap has been filled. For example, Scherer and Oshinsky (1997) studied emotional responses to synthesised melodies that were controlled according to a number of acoustic parameters, and found specifically that of all the parameters they manipulated, "tempo of the sequence seemed to be the most powerful predictor [of emotional response]"(p.339). Particularly remarkable for Scherer and Oshinsky was the robustness of tempo as a predictor. Kratus (1993) investigated the extent to which children respond emotionally to music, thereby hoping to discover the extent to which responses were merely a result of acculturation. The results of this study revealed that children as young as five years old were able consistently to judge music as happy, sad, excited or calm, and that "predictor variables of Rhythmic Activity and Articulation ... account for at least 90% of the variance in subjects' interpretation of happy and sad"(p.14). In short, not only do people relate the tempo of rhythmic patterns to degree of activity—hardly a surprising discovery, given that faster rhythmic patterns represent more activity in a very real sense—but also detected activity level has implications for the perceived emotionality of rhythmic stimuli.

That listeners perceive level of activity and relate this to emotional expression does not necessarily imply that rhythmic density actually evokes physiological reactions in the way patterns in non-musical contexts have been found to do. Assuming it does do so, however, we should expect it to cause an effect that is qualitatively different from that of loud noises, sharp timbres or even expectancy violations in patterns. Rythmic density is not a momentary phenomenon, so neither should we expect emotional response to it to be so. Given the research already cited as to the effect of repeating patterns on physiology and psychology, we should rather expect that the effect of tempo in music should be to mediate mood. A number of researchers, such as Rigg (1940), have found exactly this to be the case.[25] A

[25]Here, then, the results of empirical work that isolated tempo as an important factor in mood mediation reported in Chapter One can be seen in a broader psychological—or physiological—context.

study of considerable interest is that of Balch and Lewis (1996), which investigated context-dependent memory effects with music, a reasonably well-know effect whereby people find it more difficult to recall a series of memorised words in the context of background music which is different from music that was playing at the time of memorisation. After an initial experiment, which found that "tempo, in particular, appears to be a dimension of musical context that has significant memory consequences"(p.1357), they tested the hypothesis that tempo changes affect memory by inducing mood.[26] Participants were asked to rate their arousal in response to a variety of musical contexts which differed in timbre and tempo, with the result that faster tempi were found to lead to higher arousal; the music was unambiguously influencing mood. A further experiment, which related the two observations, led to the discovery that "during recall, more was remembered in mood contexts that matched the moods of the original presentation tempos"(p.1361). In other words, the influence of tempo on mood was significant enough to affect memory, another finding that seems rather less surprising in the light of evidence gleaned from outside of musical domains.

A glance at the repertory suggests that the relationship between tempo and mood is well known to the music industry: a clearly distinguishable rhythmic beat is a central component of much western popular music; in the case of music such as techno, rhythms are typified by an extremely fast, loud beat, whose explicit purpose is surely to excite the listener by causing physiological arousal. Sadly, none of the empirical research reviewed above provides conclusive evidence that tempo directly causes physiological arousal; the mood mediation noted by Balch and Lewis (1996), for example, might be the result of complex interactions between various factors. It seems inconceivable, however, that the powerful effect of rhythmic stimuli on physiology that has been found in non-musical domains are somehow emasculated in music. Conversely, findings concerning response to tempo in music make much more sense in the light of the more general phenomena; evidence that tempo effects physiology does not only suggest a mechanism by which tempo and mood are linked in musical context, but even predicts it.

To conclude, then, sounds and patterns are capable of causing physiological arousal and evoking emotional responses in a listener by virtue of the relationship between those sounds and patterns and the workings of our perceptual systems. There is little reason to believe

[26]This hypothesis was chosen as it is congruent with Eich(1995; see also Eich and Ryan, 2000) who claimed—and demonstrated in a variety of contexts—that context-dependent memory is mediated by mood.

that our response to sounds and patterns when they occur in music is any different from our response to them in a natural acoustic environment; in fact, all the evidence points to the converse, and is bolstered by indications that musicians deliberately harness the "natural" emotionality of sounds and patterns and subsume it into their music. However complex a musical style, however rich the musical culture, and however complex the relationship between music, a listener and the mechanisms by which emotions are evoked, raw sounds and patterns will always have certain effects on perceptual systems, and will always contribute to the evocation of emotions. Emotional response to sound and pattern constitutes an integral and fundamental part of any listening experience; it is for this reason that the first major premise of our model of emotional response is that *music is heard as sound*.

Chapter 3

Music as Utterance

"One of the most soothing things in the world is to put your tongue to the roof of your mouth right behind the teeth and sing la, la, la, la, la, la, la.*"*(Ackerman, 1990, p.205).

Whatever the role of music within a culture or society, and whatever the mode of production, a musical sound is always a human creation, and is heard as such. The production and perception of a musical sound constitutes a form of human communication; a musical sound is a human utterance. From this it follows that if the members of a particular culture—or even humankind in general—respond to human utterance in a way that is differentiable from their responses to other types of auditory stimuli, then we can expect that mode of response to apply equally to the perception of music. The first half of this chapter reviews evidence that suggests that we do indeed respond to human utterance—or more specifically to emotional expression encoded within human utterance—in a way that is qualitatively different from our responses to other types of sounds, and further that some elements of the expressive code are determined by our evolutionary history and physiology, require no learning, and are universally understood. The second half of the chapter considers the significance of this observation to a model of emotional response to music.

Communication and Vocalisation

Human evolution has provided us with an extremely sophisticated mode of communication, namely language. By manipulating our vocal folds in certain ways, we are able to produce sequences of sounds that a listener can differentiate, recognise, and from which meaning can be extracted. As with many psychological processes, the complex manipulation of vocal musculature needed to produce a series of phonemes, words and sentences is achieved without conscious awareness; likewise, a listener who has learned how to categorise the phonemes of a particular language and has developed an implicit knowledge of its words and structure will often be able to perceive the sounds and unravel the meaning without making any conscious effort to do so.[1] Despite our sophisticated verbal abilities, however, vocal communication between humans is often non-linguistic. Notably, indicators of emotional state tend to be carried either paralinguistically or by way of completely non-linguistic verbal utterances (Scherer, 1995). For example, a speaker who announces in relaxed even tones, "I don't know what I am saying" may well be simply commenting on a factual state of affairs; on the other hand, the same sentence spoken with a tremor, or shrieked in a wild, uncontrolled explosion, will be understood to have quite a different meaning. The words are the same in both cases; the differences are carried exclusively in paralinguistic elements of the vocal utterance. Someone facing an unimaginable adversary may not utter words at all, but is still very capable of communicating an emotional state unambiguously by way of a blood-curdling scream, or a terrorising shout; either of these would constitute a non-verbal utterance.

Just as we can produce or perceive a series of phonemes, we can produce or perceive non-linguistic sounds, and in either case, we need not be consciously aware of the detailed mental machinations involved. There is a crucial difference in production, however, namely that the physiological processes involved in creating linguistic sounds are under the control of the brain systems that control language production and awareness,[2] whereas the processes that give rise to paralinguistic modulations or non-linguistic utterances are under the control of emotion systems.[3] There is an important difference in the outcome of perception too: whereas the ultimate result of phoneme perception is usually conscious awareness of a concept (see e.g. Fodor, 1998), the ultimate result of perceiving a paralinguistic or non-linguistic signal is usually awareness of an emotional state (see e.g. Oatley, 1992). Whilst

[1]For an excellent discussion on the nature of subconscious perceptual processing, see LeDoux (1998).

[2]It is commonly thought that speech production, like many highly cognitive processes, is controlled by areas of the neocortex (see e.g. Damasio, 1994).

[3]There are situations in which this contention does not hold. A good actor, for example, can simulate emotional utterances at will with considerable success. As Pinker (1997, p.415) comments, however, it is extremely difficult successfully to fake an emotional expression or any sort: "People don't really believe that the grinning flight attendant is happy to see them".

these distinctions might seem intuitively obvious, they do not provide much solace for the emotion researcher interested in paralinguistic communication: linguistic signals consist of discrete sounds, all of which are acoustically differentiable by a small number of parameters,[4] but there is no such clear pattern for the paralinguistic and non-linguistic sounds that communicate emotional state. It is rather difficult systematically to investigate our ability to detect and respond to non-linguistic emotional utterances when it is not possible to isolate those elements to which people are meant to be responding![5]

Despite the problems, scientists over the last few decades have carried out significant research in this area, most of which has focused on isolating regularities in the acoustical properties of emotional utterances. This research has normally taken one of two approaches. One approach has been to take recordings of emotional utterances—genuine or simulated—for which the emotion portrayed is known, and to subject the samples to extensive acoustic analysis. The other has been to study vocal physiology to see how it changes when communicating different emotions, and then to consider what the acoustic correlates of sound production through that physiology might be.

Research in the first category has tended to reach a broad consensus on the basic principle that acoustical cues in vocal signals can unambiguously encode reasonably specific emotions; on the other hand, there is rather less agreement on what aspects of the acoustic signal play the most important role in determining the nature of the emotion encoded.[6] Kappas, Hess, and Scherer (1991) provide an excellent synthesis of research in the field, in which they point to the importance of intensity, intonation and fundamental phonation frequency, the most important parameter being fundamental phonation frequency. Their findings may be summarised as follows: boredom is characterised by low fundamental frequency and a low intensity; irritation leads to an increase in fundamental frequency and intensity, coupled with downward intonation contours; sadness and dejection lower the fundamental frequency significantly; by contrast, fear evokes a vast increase in fundamental frequency.

[4]Speech sounds are categorised as phonemes, of which there are approximately 40 in the English language (see e.g. Moore, 1989).

[5]This problem parallels that raised by Sloboda (1992) in his discussion of emotional response to music, where he notes that "little is known about precisely what it is in the music that they [people] are responding to"(p.39). The difference here, however, is that the problem is rather more narrowly defined.

[6]A possible rationale for this lack of agreement can be found in Scherer (1992, p.44), which blames methodological difficulties of systematic study of vocal affect expression in animals on issues such as "the fact that vocalisations studied in the field are often described only by the use of more or less appropriate verbal labels, sometimes supplemented by spectrographic illustrations".

Whilst Kappas et al. present an impressive array of data, it should be noted that the differences in description between the various mappings is far from unambiguous. A notable point of ambiguity, for example, concerns the lack of a properly defined base-line around which the parameters allegedly vary. Further, researchers such as Scherer and Oshinsky (1997) have found that other aspects of the acoustic signal, such as harmonic richness, can be just as important as fundamental phonation frequency in determining emotional state, and an additional complication, noted by Tischer (1991), is that the specific role played by various acoustical parameters in the communication of vocal emotion varies dramatically depending on temporal location within an utterance. A concise summary of this work can be found in Scherer (1992).

Supporting evidence comes from research on the acoustic parameters involved in the portrayal of emotions by stage actors. For example, Williams and Stevens (1972) employed a long-term-spectrum analysis technique to analyse signals of actors' voices whilst reading a play; such analysis offers a standard method for examining the timbral quality of an acoustic signal, as it gives an indication of the number of upper partials present in the signal and a quantification of their relative strengths. The researchers found that the relative strength of various partials in a vocal signal varied significantly with the emotional mood of the text: anger evoked a strong presence of partials in the supra-1kHz range; by contrast, sadness evoked almost none. Despite these promising results, relatively little research on emotional communication through vocalisation has made systematic use of long-term-average spectral analysis. Aures (1985), Malloch (2000) and Pollard and Janson (1982) have each suggested sophisticated methods for analysing the timbral qualities of a vocal signal that allow for quantitative comparison to be made along naturally-perceived dimensions such as "roughness" and "width", comparisons that could well be of value here. The power of such techniques has been demonstrated amply by Malloch (2000), in his investigation of mother-infant interactions, but they do not yet seem to have been applied generally to the acoustic analysis of emotive speech.

The second approach to investigating the relationship between emotion and vocal signals has been to examine vocal physiology and the way in which changes in general physiology brought about by a particular emotional state affect sound production. The rationale for this approach has been stated succinctly by Sundberg (1987, p.154), who comments that "it does not seem farfetched to assume that emotions and attitudes have articulatory effects". Likewise, Scherer (1995, p.240) has stated that "given the manifold determinants of voice production processes, even slight changes in physiological regulation will produce variations in the acoustic patterns of the speech waveform"(p.240). An abundance of evidence demonstrating visibly differentiated physiological correlates of certain emotional states (see e.g. Wagner, 1989) indicates that much of the variation in fundamental phonation frequency, timbral quality, tempo and articulation that occurs in the acoustic signal of emotional voices could indeed be attributed to the effect of ANS arousal on the vocal system. In other words, the acoustic qualities of a voice will vary depending on the emotional state of the voice's owner, owing to changes in physiology causes by mechanisms that operate outside of conscious control.

The fact that researchers have found a correlation between vocalised emotion and certain aspects of the acoustic signal, and that these correlations can be attributed to the effect of arousal on vocal physiology, does not in itself imply a paralinguistic communication system; communication would require that emotion encoded within a signal be at least perceived, and at best responded to by a listener. There is a wealth of evidence, however, to corroborate the suggestion that we can indeed perceive differences in acoustic signals caused by arousal-evoked physiological changes, and can judge from these signals the emotion being communicated. A plethora of independent studies asking listeners to judge the emotional quality of speech samples has found, without fail, that participants can do so with a high degree of consistency for a wide range of emotional expressions. In other words, those sounds whose acoustic profiles were produced by utterers in a given emotive state are also perceived as portraying that state; emotion encoded in an acoustic signal really is communicated.

All these findings are encapsulated in an observation made by Kappas et al. (1991) which notes that "from vocal cues, listeners can ... determine affective states when presented with only a few sentences or, even more clearly, shouts or cries"(p.218); a thorough review of research that might lead to such a conclusion can be found in Scherer (1981, 1986). Sedlacek and Sychra (1963) have noted that listeners can judge the emotive content of a speech even if it is spoken in a language with which they are completely unfamiliar. The implication here is that the production and perception of emotive cues in vocal utterance is certainly not merely a phenomenon deriving from linguistic convention; rather, it may well form part of a universally human signalling system that operates independently of culture-specific knowledge. Just as we respond innately to certain types of sound (see Chapter Two), perhaps we also respond innately to emotional signals encoded within human utterances. It is to this proposition that we now turn.

Utterance, Evolution and Development

The findings of Sedlacek and Sychra (1963) would suggest that even if our ability to perceive emotional cues encoded within vocal utterances is not innate, there is at the very least a surprising degree of cross-cultural consensus as to what emotional content is being conveyed by a vocal signal. A similar implication comes from Frick (1985) in his cross-cultural study of vocalised emotions, which indicated that many emotions and their vocalisations exist universally.[7] On the other hand, Scherer (1995, p.240) reminds us that "vocalisation had developed in part as a social communicative signalling system, [in which] the simple externalisation of internal states has been supplanted by the display mechanism producing a specific impression in the listener". Although this display mechanism may well be a part of the human phenotype, specific displays could be highly connected with a particular society or culture.[8] To the extent that perception of emotional signals encoded within vocal utterances is a function of cultural circumstance, it is no different from any other cognitively and socially mediated phenomenon, and deserves no special mention here.[9] If from these signals, however, we can perceive the emotional state of an utterer owing to our status as humans, regardless of cultural context, then surely such sounds deserve a special place in any account of emotional response to auditory signals, including music.

Two very different types of research have provided persuasive evidence that the production and perception of emotive cues in vocal signals are not only innate human abilities, but are also ontogenetically early developments. The first of these stems from Darwinian naturalism; the second from developmental psychology. Darwin (1872) claimed that in order

[7] The extent to which emotions are expressed with similar patterns cross-culturally has been an issue of considerable interest to emotion theorists in general. Ekman (1992, 1993, 1994b) provides compelling evidence that perception of and response to facial expressions takes on an identical form pan-culturally.

[8] Examples of culture-specific emotion displays abound. Public displays of grief in many cultures are accompanied by a type of wailing that is highly ritualised and culturally-determined; instances of such wailing have been the subject of a number of anthropological studies (e.g. Urban, 1988).

[9] The importance of cultural context for the interpretation of human communicative signals is discussed extensively by Sperber and Wilson (1995); it is also considered in Chapter Four.

to understand human non-verbal communication, such as shouts, screams and wailing, we should study vocal communication in the animal kingdom. His contention was that the processes involved in producing animal and human vocalisations were phylogenetically continuous, the implication being that the exclusively human ability to communicate verbally was a much later evolutionary development than the ability to make communicative sounds. Given that many different species of animals possess voices, and that most can communicate in some way using these voices, it seems highly plausible that there are common modes of usage. Even so, rather remarkable is recent evidence that despite differences in vocal systems of various animals, there is a high degree of similarity in the general acoustic patterns various animals produce in the same situations (Tembrock, 1975; Morton, 1977). For example, animals who are being threatened—presumably thereby experiencing fear—tend to produce sounds with a wide energy spectrum and high onset amplitude just like a scream; on the other hand, regular repetitions and low energy "contact calls" tend to typify relaxation and contentment. Interpretation of animal vocalisation requires a degree of caution, as it is not necessarily possible correctly to infer the cause or "intention", but this evidence is nonetheless most compelling. It is quite possible that some of the cross-species similarities that have been found in the animal kingdom have a rather mundane origin, for example, "the physical qualities of sounds may play important roles in communication, because some qualities may be more suitable for communication across large distances whereas other qualities may allow rapid and fine patterning in proximal communication" (Papoušek, 1996, p.40). Alternatively, similarity may be a function of similar physiological means of production. Despite these reservations, however, it is highly likely that "the adaptive significance of vocal communication has resulted not only in innate anatomical predispositions but also in a strong intrinsic motivation for its use in primates, including humans"(Papoušek, 1996, p.41). One could add to this that in a world in which many species share the same terrain, an ability to communicate across species boundaries might have significant adaptive advantage!

Although the above observations serve as a persuasive argument that acoustic signals shaped as a result of human vocalisation constitute some form of "absolute" communication system that transcends cultural or even species boundaries, they tell us nothing about the ability of such signals to *evoke* emotion in the perceiver. Whilst one might expect that cognitive appraisal of an emotive signal could lead to induction of emotion in the perceiver through a process such as empathy (see e.g. Eisenberg and Strayer, 1986), it would be enlightening to know whether induction of emotion through perception of emotive human utterances could be achieved as a result of innate human disposition; certainly such knowledge would have implications for the study of emotional response to utterances in music! Recent research in developmental psychology sheds some light on this issue by showing that despite a limited ability to vocalise, even newborn babies are perfectly capable of recognising *and being affected by* human utterances. Not only do they respond to utterances in a way that is different from their responses to other sounds, but it seems that they are also readily able to detect and *respond to* emotional cues embedded within vocal utterances.

Papaeliou and Trevarthan (1998) have noticed that newborns appear to prefer listening to human speech than any other sound, regardless of what language is being spoken, a finding that parallels that of Sedlacek and Sychra (1963). Given the significant differences between the sounds of diverse languages, it seems that whatever the babies are attracted to, it must be something distinctly human about the vocal utterances that differentiates them from other sorts of sounds. A complementary finding concerns "motherese", the nonsense sounds that mothers use when communicating with pre-lingual babies: one might expect that such sounds would be related in some way to the mothers' native tongue; it seems, however, that mothers who speak languages as diverse as American English and Mandarin utter similar motherese to their babies (Greiser and Kuhl, 1988). A possible explanation for this is that mothers are producing—probably non-consciously—the sorts of vocalisations that their babies will innately understand to be comforting and affectionate, or to take an evolutionary stance, mothers are producing vocalisations that have over an extensive period of time have had proven success at soothing infants. It is telling that mothers who are suffering from depression find it difficult to communicate with their babies, and that the babies, in their turn, fail to respond to the sounds that their depressed mothers do make (Robb, 2000; see also Gratier, 2000).[10]

From a somewhat different perspective, Scherer (1992) suggests a mechanism whereby emotion encoded in vocalisation might not only be perceived but actually induced in a listener. After a notion first suggested by Lipps (1913), Scherer notes that much psychological work on facial expressions of emotion (e.g. Buck, 1980) has pointed to an apparently innate

[10]Trevarthen (1993) found a similar effect with singing: depressed mothers find it difficult to sing to their babies, and the babies, in turn, rarely respond to such singing.

human propensity to mimic:

> "The automatic tendency toward motor mimicry or observed expressive behavior serves as the mechanism for emotion transfer. In other words, the observation of someone else's expressed emotion will, at least in a rudimentary fashion, evoke the same emotion is us, the observer"(Scherer, 1992, pp.55-56).

He suggests that a similar effect might exist in the domain of vocalisation: we could perceive a vocal signal as a physical expression and, through this automatic process of mimicry, we would feel that expression; hence the emotion would be induced. Although Scherer takes pains to point out that this is merely a "speculative hypothesis"(p.57), it perhaps does not seem so farfetched.

Evidence from animal studies and developmental psychology, then, along with a resurgence of interest in evolutionary explanations in general,[11] has led to a general consensus that the ability to produce and perceive non-verbal and paralinguistic emotional utterances is innate and that the relationship between emotion uttered and acoustic signal is somehow absolutely determined on some level. This does not, however, preclude the possibility that hard-wired innate abilities to perceive certain types of sound as emotional utterances could be supplemented by culture-specific learning. In fact, just as we learn to recognise the phonemes of our own language and the meanings of the words formed by them, there is no reason to believe that we cannot in addition learn a set of non-verbal utterances, or at least paralinguistic signals, and their associated meanings.[12] An example of learned paralinguistic signals that are specific to, say, the English language, are sarcastic tones of voice: everyone can produce one, and everyone knows what it means. Such learned meanings, however, are nothing more than associations, and to recognise one is no different from recognising a word or phrase in a language; they are qualitatively quite different from the phenomenon under study here. There is, however, something unique about a "natural" human utterance that makes it totally different from any other kind of sound, and that is fundamental to the communication of emotions: emotion coded within a human utterance will perceived and *can be responded to* by any human listener.

[11]The huge popularity of authors such as Stephen Jay Gould and Richard Dawkins stand testament to this interest, at least in the UK!

[12]Such learning could take a form similar to that noted in Burt et al. (1995), where participants in an experiment were able to learn to associate specific sounds with levels of urgency in a warning signal.

Music as Utterance

If the perception of emotive cues encoded within vocalisation is a universal phenomenon, and the acoustic profiles of those cues is consistent pan-culturally, we should expect that any musical sound that shares an acoustic profile with such an utterance also has the potential to carry the relevant emotional signal. We should expect further that signals communicated by such musical sounds would be perceived and could be responded to in a way that is no different from perception of and response to any other utterance; in other words, emotional response to such music would be, at least in part, emotional response to the apparent human utterance. In many cultures and societies—including our own—one of music's primary roles has been to communicate emotion; it is here claimed that it owes much of its success to the fact that such music either consists literally of vocal expression, or else emulates vocal expression by means of analogous sounds. The parameters along which musical sounds vary are the same as those along which emotional cues in vocalisations vary; the building block of music are those very same units that form the basis of non-verbal communication.

Many evolution theorists (e.g. Vaneechoutte and Skoyles, 1998) have assumed that music, like language, evolved out of primitive vocalisations, and that those vocalisations had originally served to communicate basic emotional states, such as fear or contentment. As music has developed, it has remained tightly coupled to the voice and vocalised affect. Vocal music is central to virtually all musical cultures, and in some, such as in Iraqi *Maqam* or Spanish *Cante Jondo*, the quality of the voice has an almost exclusive monopoly on the emotional tenor of a performance. In the western art music tradition, a similar story emerges: composers have often been preoccupied with the communication of emotion, and whenever they have, the debate seems to have focused on vocal expression. Possibly the most obvious example of this is the dramatic style change that appeared in Italy at the start of the sixteenth century: one of its major exponents, Caccini, declared in his introduction to *Le Nuove Musiche* that music must honour the natural contours of vocal expression; he was deeply critical of the highly embellished style of the day, which be referred to "as a kind of tickling of the ears of those who hardly understand what affective singing really is .... nothing [is] more inimical to affective expression"(p.47). Similar sentiments have been reflected by other composers since, notably by Rousseau.[13] Presumably, if music has been written deliberately to promote

[13] See Verba (1993) for an excellent discussion of the French Enlightenment musical debate.

affective expression of the voice, and has been successful in its quest, any emotional signals encoded by the singer will be communicated successfully to the listener, just as they have been shown to do in non-musical contexts.

In his extensive study of western art music, Cooke (1959) has illustrated that composers over the centuries have perpetually used the same emotive musical phrases in their music despite dramatic changes of various musical styles. Whilst acknowledging the possibility that convention has an important role to play, he assumes that there must be something fundamental about these musical patterns that seem persistently to re-emerge:

> "It would be useless to deny that the continuous and consistent use of certain terms of musical language throughout five centuries or more must have conditioned us to accept them without question .... But it is difficult to believe that there is no more to it than that."(p.25)

Cooke tried to explain the phenomenon in terms of the harmonic series, but in the light of the universality of vocalisation in music, and the observations offered by theorists such as Vaneechoutte and Skoyles (1998), it seems much more likely that these musical patterns are abstractions and formalisations of natural vocal expression. Cooke himself noted that an analogy between music and language "can best be understood on a primitive level"(p.26), citing as an example a groan, which becomes a formalized sung "ah!", which could possibly even be abstracted further to become a pattern of a flattened sixth falling to a third. Equally, Meyer (1956, p.260) noted the implicit connection between musical patterns and affective vocalisations when he stated that "because moods and sentiments attain their most precise articulation through vocal inflection, it is possible for music to imitate the sounds of emotional behaviour with some precision".

Music of some kind exists in every known culture (Blacking, 1995). Further, within a particular culture or sub-culture, it always has a specific role—or range of roles—whether ritualistic or simply social (e.g. Gregory, 1997). A role-congruent mode of listening must by definition be culture specific; the hearing of emotional cues, if the culture dictates that such cues should even exist, is likely to be a highly cognitive and metaphorical activity. This observation might encourage someone wishing to counter Meyer's claim to argue that moods and sentiments may "attain their most precise articulation through vocal inflection" but that music does not actually "imitate the sounds of emotional behaviour" with any precision at all; the link between natural vocalised expression and its implementation in music is merely an assumed, culturally-contingent metaphor. If this were found to be true, and if culture-specific metaphor provided the only connection between music and vocal utterance, the findings would have significant impact on the present thesis: the special status afforded to utterance here would hardly seem appropriate! Given the overwhelming evidence that emotion encoded within vocal expressions transcends cultural boundaries, however, an exclusively metaphorical link seems highly implausible. In addition, even within the realm of music, there exists plenty of evidence to suggest that response to certain genres transcends cultural variation; there also exist cross-cultural similarities that can only satisfactorily be explained in terms of vocal music containing emotional cues that are perceived as human universals.

One of the most striking examples is provided by Trehub, Unyk, and Trainor (1993), who carried out a fascinating study of a musical form that is found ubiquitously across cultures, namely the lullaby. They found that people are able to differentiate lullaby melodies from non-lullabies with ease regardless of the musical culture from which they originated, and that all lullabies use similar patterns. All are typified by repetitive motifs and slow-moving contours; they tend to be sung in a quiet, gentle voice, presumably not rich in high harmonics. In other words, they create an acoustic signal that shares the very same patterns as those that typify vocal sounds—both human and animal—that have been shown to signify contentment and relaxation (see earler in this chapter). Given that the patterns are so ubiquitous, and that lullabies have a clear and well defined function, it seems reasonable to assume that the two facts are related: lullabies use the patterns that they do because these patterns are perceived and responded to as if they were soothing human utterances. Unfortunately, little research seems to have investigated whether cross-cultural similarities extend to other musical genres.[14]

Despite the lack of research in the area, we might expect that cross-cultural similarities found in lullabies would, in fact, extend to other genres of music, owing to a far more fun-

[14] An obvious reason for this dearth of research is that the differences between the roles that music plays in various cultures often precludes easy cross-cultural comparisons. Scherer (1992, p.44) made a parallel point in his discussion of the methodological difficulties associated with investigating animal vocalisation: it is extremely difficult to compare vocal affect signals across species because "calls are generally studied within the confines of the particular behavioral repertoire of the respective species".

damental, pan-cultural link that has been found between musical behaviour and human utterance. This evidence can be found by looking once again at the interaction between mothers and their babies. Papaeliou and Trevarthan (1998, p.20) note that "it is striking that musical play spontaneously emerges as a key component of how most parents everywhere communicate with infants ... a few months after birth". What is perhaps more striking is that this musical play takes a very similar form pan-culturally. Fernald (1992) commented that the distinguishing features of "motherese" are its mannered wide intonation and its clear stress patterns; in other words, what distinguishes motherese from normal language—or from highly emotional non-verbal utterances, for that matter—is its musicality. As Fernald (1992, p.278) put it:

> "In both soothing and arousing the infant, the mother accomplishes her intentions, in part, through her use of sound. However, the communicative force of her vocalization derives not from their arbitrary, assigned meanings in a symbolic code but from the more immediate power of music to alert, to alarm, to sooth, and to delight".

The musical nature of motherese gives us a possible explanation for the Greiser and Kuhl (1988) findings, cited earlier, about its cross-cultural similarity that transcends native language: language is simply not at issue here; instead, mothers are engaging with their babies using a more basic form of communication in both evolutionary and developmental terms, namely, the music of the human voice. Music is playing a communicative role that is independent of a particular culture or society. So overwhelming is the volume of evidence that these so-called musical sounds are fundamental to the interaction between mother and baby that Malloch (2000, p.47) declared:

> "It is our contention that the ability to act musically underlies and supports human companionship; that the elements of Communicative Musicality are necessary for joint human expressiveness to arise, and lies beneath, to a greater or lesser extent, all human communication".

It is very telling that "music" is the word used by so many infant psychologists to refer to this interplay between mother and baby. Although the interaction is described as "musical", the sounds produced are surely not "music" in the usual sense. The allusion being made is simply that the sounds produced are very much *like* music. The simile, however, is somewhat misplaced: what the interplay between mother and baby is actually very much like is positively valenced, emotionally-charged vocal utterance; it is referred to as musical because *many musical sounds are very much like the vocal utterances that humans have an innate ability to produce and perceive*. Moreover, parameters along which music operates are the same as those which encode emotional cues in acoustic signals and are perceived and quite probably responded to as such.

If human vocal utterances are able to express emotions that we perceive, and if there are grounds for believing that expressed emotions may even be induced in the perceiver, it seems plausible that any music that *sounds like* human vocal utterance should be just as expressive, whether Maqam, Montiverdi or Motherese. Even purely instrumental music, which has no vocal element at all, could be considered to be human utterance if those mechanisms responsible for our perception of emotional cues in vocalisation still detect those cues in the music.

Research on the perception of emotional expression in musical performance sheds considerable light on this issue by indicating that sounds produced by musical instruments are indeed heard as human utterances. In one particularly revealing study by Gabrielsson and Juslin (1996), the acoustical properties of a whole range of performances on different instruments were examined in order to elucidate "expression profiles" (p.85) for specific emotions. The study showed that, regardless of instrument, there was a core set of expressive features that typified various emotions: anger was typified by quick tempo, high intensity, marked contrasts between long and short notes, harsh timbre and distortion; by contrast, sadness was normally expressed by slow tempi, a low intensity sound lacking in high harmonics and a slow vibrato; fear was mainly characterised by irregularity of timing, very fast vibrato and variations in sound intensity. Although they did not put it this way themselves, Gabrielsson and Juslin (1996) found that the sorts of acoustic signals that characterise instrumental performances that are intended to express emotional states—and are perceived to be expressing these—do indeed display the same fundamental characteristics as the equivalent non-verbal human utterances.

This research raises two important issues in terms of our present concerns. The first, already stated, is that the parameters involved in producing and presumably perceiving emotion in musical performances are the same as those involved in the production and perception of emotion in human utterance.[15] The second important issue concerns the relationship between vocal and instrumental music. Much of the evidence presented here concerning music as human utterance has concentrated on vocal "music", whether lullabies or motherese. This could easily lead to the criticism that the link pointed to is merely that between speech and song, hardly a surprising discovery since both are forms of vocalisation, which is human utterance by definition. Gabrielsson and Juslin (1996), on the other hand, worked with instrumental musicians and yet still found similar patterns of results, indicating that the production and perception of emotional state in non-vocal music involves the same acoustic signals—and therefore probably the same mental processes—as those involved in vocalisation. In other words, a musical sound that carries emotional cues is a musical sound that behaves as though it were a human voice. The prominence of non-vocal music is a uniquely western phenomenon—few other cultures in the world have such a developed purely instrumental musical tradition—but the present evidence suggests that even emotion response to this peculiar abberation can be seen, at least on one level, just as a specific instance of a universal human phenomenon, namely the ability to perceive and respond emotionally to human utterance.

We have an innate ability to produce vocal sounds which encapsulate an emotional state; we have an equivalent ability to perceive such signals in an acoustic waveform and therefore respond to them in a way that is differentiated from our responses to other sorts of sounds. Both of these processes can occur outside of conscious awareness. The history and probably the pre-history of music is inextricably tied with vocalisation, the principle mode of human emotional communication, and musical performers attempting emotional performances tend to produce sounds whose acoustic profiles—in terms of timbral and dynamic contour—match those of emotive vocal utterances. In addition, a wish for music to be expressive has led, at least in western art music, to melodic styles that shadow the general

[15] A major difference, of course, is that a musical performance is always a deliberate act whereas a vocalisation need not be, but this difference does not undermine the hypothesis because a vocalisation is very often a deliberate act, and when it is, its communicative power is not lessened. A shout of anger, for example, may very well be a deliberate act, but the acoustic profile of the sound uttered will nonetheless depend largely on physiological factors which are not under conscious control, and could carry a powerful emotional message (but see also Pinker, 1997, cited earler). It would be interesting to know whether musicians are normally consciously aware of the acoustic parameters that they manipulate in order to achieve the desired emotional effect; sadly, this is not a research question that appears to have attracted any interest.

contours of natural vocal utterances, albeit in a distilled and formalized way. The result is that we hear musical sounds as human sounds and respond to them as though they are vocal utterances; it is this that leads to the conclusion that musical sounds *are* human utterances.

This conclusion has implications for research on emotion in music, not least because it suggests that so-called intrinsic sources of emotion in music, such as melodic contour, in fact derive their communicative power from a phenomenon that lies outside of the musical domain, something that one might in fact call *extrinsic*! It also suggests that analysis of expressive musical performance in terms of the extensive research that has been carried out on the emotional expressivity of vocalisations might be highly enlightening.

Utterance, Music and Gesture

Whatever the processes that enable us to produce and perceive human utterance, there is evidence to suggest that they are not domain specific, that is, their operation is not necessarily restricted to the auditory realm. We have already seen that processes whose evolution is inextricably tied with vocalisation can be applied to the production and perception of non-vocal musical sounds, and Scherer (1992) has pointed to the plausibility of a link between vocalisation and induction of emotion through mimicry of perceived physical expression. In addition, there is a strong link between human utterance and bodily motion, or gesture. Rimè and Schiaratura (1991), amongst others, have noted that human speech is invariably accompanied by gestural movement, and further that gestures are an indispensable part of the communication process. Of particular interest to Rimè and Schiaratura is the fact that even though humans are exposed to gestures all the time, they rarely consciously notice them; the researchers state without irony that "with the gestures and bodily movements of the speaking person, we are faced with embodied thinking"(p.241).[16] Often, if any specific aspect of communication is expressed by gesture, it will be the communicator's emotional state, because physiological concommitants of emotions are known to be coupled with motor activity (see e.g. Thayer, 1989).

Gesture plays an important role in music too. Musical discourse is littered with metaphors

[16]Such an assertion is somewhat reminiscent of Lakoff (1987, p.xv) when he writes that "human reason ... grows out of the nature of the organism and all that contributes to its individual and collective experience: its genetic inheritance, the nature of the environment it live in, the way it functions in that environment ... and the like".

of motion, such as "up" and "down", "fast" and "slow", or the constant dynamic between tension and relaxation that forms the basis of western tonal music. Musical performers frequently mimic musical shapes with body movement, and can even be unaware of the extent to which they are doing so (Davidson, 1991);[17] in many cultures, including our own, music and dance often exist together as a single entity, as has been noted by Krumhansl (1997a).

This relationship is not purely metaphorical, but appears to have developmentally early roots, as can be seen by observing for one final time the musical play between a mother and her infant, this time focusing on the role of movement: frequently, the mother will vocalise sounds that mimic the dynamic contour of an infant's actions (Papaeliou and Trevarthan, 1998); when she is singing, she will look for sympathetic gestural movement from the infant, such as movement of eyes and limbs; the infant will rarely disappoint. Sundberg (1987, p.157) provides a different, but complementary perspective when he shrewdly observes that as we perceive acoustic signals emanating from a human voice "we project the voice timbre we hear from our own voice organ, and we describe the timbre via this imagined phonation". In other words, we perceive the timbre of an utterance in terms of the physiological motion required to produce it; we perceive an utterance as a gesture.[18] Such a notion provides another basis on which instrumental music might be perceived as human utterance: performing a musical instrument involves physical, gestural activity just as does singing; perhaps it is this gesture that is encoded within an instrumental "utterance", and is perceived as such by a listener.

No discussion of gesture and music would be complete without reference to the Clynes (1977) theory of *sentics*, so we will turn to this theory now. According to Clynes, emotion is communicated through dynamic forms that come into being when, for example, music is performed. These forms, which he terms *essentic forms*, are particular shapes that embody particular emotions. Evidence for the existence of essentic forms comes from a series of experiments in which participants were given the task of matching patterns of finger motion to sets of emotion words; it was found that they were able to do so with a great degree of consistency, whether they had learned to perform the finger actions themselves—through imitation—or if they merely watched them on a television screen. Clynes and Nettheim

[17]Cited in Clarke (1993).

[18]Sundberg's observation nicely balances research (cited earlier) indicating that the nature of an emotionally charged vocalisation could be determined by physiological effects of ANS arousal, and lends support to Scherer's (1992) hypothesis.

(1982) devised a way of translating the finger movements into sound patterns, using a series of pressure and motion sensors connected to an analogue synthesizer, and then played these to a number of western and non-western participants. Apparently the participants were able to match the sounds to emotion categories with a high degree of accuracy, regardless of their cultural background.[19]

Although Clynes's theory has not found a central place in the canon of music psychology research, in the present context its tenets do not seem too outlandish. Clynes and Nettheim note that despite the lack of precision with which dynamic markings are indicated on a score, and despite musicians' inability to play at a precisely prescribed dynamic level, any good musician will precisely alter the amplitude variation between notes in order to achieve the desired emotional effect. It is claimed that this ability is the result of fitting the individual notes into some sort of emotionally-charged "concept":

> "The concept springs from the total empathic knowledge of the musician ... and guides the unfolding of the chain [of individual notes]. It is thus also that the listener becomes more aware of a missing link in a faulty performance – sound without essentic form"(p.51)

If the nebulous notion of "concept" is replaced with "gesture", or perhaps even "utterance", the statement makes a lot of sense: when a musician performs, an overall conception of emotion is all that is required to produce the relevant acoustic cues; if these cues are not produced, they cannot be perceived by a listener, so the performance will be devoid of human utterance. The fragmentary sounds produced from finger motion in the experiment discussed above hardly constitute real music, so it might be considered a little premature for the experimenters to have drawn such sweeping generalisations out of them. However, individual utterances are by definition short and often fragmentary. If the spectrum profiles and patterns of fundamental frequency variation were comparable to those found in human or animal vocalisations—not an unlikely possibility given the close link between physical gesture and utterance—then there is no reason why they could not have been perceived as such.

[19] A number of methodological problems with the experiment warn against over-interpretation of these results. One particularly glaring issue was the considerable flexibility that the experimenters gave themselves when converting finger-motion patterns into sounds; for each essentic form, they altered parameters that affected the way finger motion was converted, so that each sound was subjectively appropriate to the form it was meant to represent.

When Clynes and Nettheim (1982, p.51) declare that "it is difficult to remain unaffected in the presence of a true, authentic expression of grief, or of joy" they are referring to the power of so-called essentic forms. In the light of the evidence presented in this chapter, it might be more appropriate to declare that it is difficult to remain unaffected in the presence of an emotional human utterance. We perceive the emotional cues and have an implicit understanding that they been produced non-consciously as a result of the emotional state of the utterer. There may be many modes of listening to music, but one of these always involves the unquestioned assumption that music is human utterance. Hence, that music is heard as utterance, and that hearing it as such is essential to the emotional nature of musical stimuli, is the second main premise of this thesis.

Chapter 4

Music as Context

> *"The musical object is never isolated, any more than are its listeners or its producers"*(Feld, 1984, p.6).

We have now seen how the sounds of music can cause physiological arousal in a listener simply because they are auditory stimuli; sometimes, they can communicate powerful emotive signals, merely by virtue of their similarity to human utterances. Yet, as has already been noted, we do not run screaming from the concert hall when the canons fire in Tchaikovsky's *1812 Overture*; we do not cry in anguish at our inability to provide aid when we hear the strains of *Cante Jondo*. The reason? Cognitive mediation tells us that the loud bangs should not be a cause for alarm, and that the anguished strains of the *Flamenco* singer are part of a deliberate performing act. We hear the sounds and the utterances as part of a musical experience; in short, we hear them in context.

Just as our understanding of the ramifications of loud sounds and anguished utterances is mediated by a musical context, so is music itself—as the epigraph to this chapter implies—understood and responded to within a tangled web of interdependent contexts. A listener's mood, the situation in which music is heard, past experiences associated with a particular piece or style of music, coupled with knowledge of the circumstances of composition or performance, its political or social significance and any other extraneous but significant information constitute the context in which the listener hears, understands and responds to the music; these provide a framework within physiological arousal from the sound and utterance domains can be interpreted and, as will be seen, are in themselves crucial sources of emotion. Discussing the importance of context for the interpretation of spoken utterance, Sperber and Wilson (1995, pp.15-16) write:

> "A context is a psychological construct, a subset of the hearer's assumptions about the world. It is these assumptions, of course, rather than the actual state of the world, that affect the interpretation of an utterance. A context in this sense is not limited to information about the immediate physical environment or the immediately preceding utterances: expectations about the future, scientific hypotheses or religious beliefs, anecdotal memories, general cultural assumptions, beliefs about the mental state of the speaker, may all play a role in the interpretation".

This definition, surely, applies equally well to the experience of listening to music.

This chapter considers the relationship between context and emotional response to music, and seeks to demonstrate how a vast proportion of reported phenomena involving emotion and music can be understood in terms of context-induced emotion. The overriding claim will be that to listen to music is, at least in part, to hear sounds, utterances and even musical structure itself in and as context.

A Context for Listening to Music

Clear evidence that responding to music is a contextual phenomenon comes from the apparently trivial observation that for many people, to listen to music is to remember past thoughts, events and experiences. A song, a symphony or even just a musical style, once heard, is not remembered in the abstract but in the context of its hearing. Music is associated in memory with a plethora of thoughts, events and feelings; thus to re-hear music is to evoke a whole web of related associations.[1] Further, it would appear that these contextual phenomena are important sources of emotion evocation, as is illustrated by a large-scale survey carried out by Gabrielsson and Lindström (2000), in which thousands of respondents were asked to submit reports on "strong" emotions that they had experienced in response

[1] "Association" is a rather loaded term in psychology, referring as it does to various schools of thought and theories of mind. No such allusions, however, are intended; the word is used in its common-usage sense to refer to concepts or thoughts that are connected together in some way. As Crowder, Serafine, and Repp (1990, p.473) declared: "The term *association*, by itself, may connote many things theoretically, such as rote learning, Pavlovian conditioning, or the antediluvian mists of precognitive psychology. However, the term is theoretically empty; it simply connotes an experimental fact, that Events A and B stand in a particular empirical relationship because of their history of co-occurence".

to music. Respondents gave details about the music in question, the circumstances of listening and as much relevant detail as they could remember; analysis of the data revealed that whilst some emotional experiences involving music were attributed directly to the music itself, others were caused by the circumstances of listening, the triggering of a memory, or general state of mind. Crucially, the researchers noted that "although one may be tempted to believe that music is the dominant factor [in evoking emotions], this is often not true"(p.107). A similar story is told in an investigation by Baumgartner (1992): almost three quarters of the respondents to a questionnaire reported having been moved by particular pieces of music owing to autobiographical experiences associated with those pieces; for two thirds of these respondents, music was associated with a romantic involvement. Such findings should not be unexpected: people always remember salient experiences better than they do information of no direct personal relevance (Bower, 1981; Strongman and Russell, 1986); if music forms part of a salient experience, associations between that music and other aspects of the experience are bound to be stronger than those in which music is part of non-salient, quintessentially forgettable ephemera.

To suggest that familiar music can evoke autobiographical memories in a listener does not necessarily imply that the memory, once triggered, becomes a salient context for the listening experience itself; instead, it is quite possible that it could evolve a life of its own quite separate from the music. There exists convincing evidence, however, that autobiographical memories sparked by music do in fact remain central to a listening experience. An example of such evidence is provided by Waterman (1996). He carried out an experiment in which participants were played a set of musical extracts, and were asked to click a button if they felt moved at any point during the playing; anyone who had clicked the button during an extract was asked to declare why he or she had done so. In many cases, the rational given was that the music sparked a memory of some salient event; in other words, it was not the music but an associated memory that caused the emotional response. At first sight this finding seems rather unremarkable, merely confirming that music can indeed spark associated memories and that these in turn can have emotive connotations. Of particular significance here, however, was the exact nature of the task: by asking listeners to click a button at any *point* during the music, Waterman was focusing participants' attentions on dynamic, moment-by-moment aspects of the music,[2] but even with their attention thus focused, a primary mode of response involved the evocation of memories for past events and past experience. Waterman's findings are not unique; Lowis (1998) found very similar patterns of results using a similar experimental situation; specifically, 65% of the participants in his study were reminded of a past event or experience during the playing of familiar musical passages. In short, autobiographical memories, once triggered, impinge directly on emotional response to music itself; they really do constitute a salient context.

In the western world, music is omni-present, and is so often played in mundane situations—sounding in locations such as supermarkets, cars and bus-stops—that most people must constantly be involved in scenes in which it plays a part. It is hardly surprising that music often serves as a memory trigger: on the basis of probability alone, we could predict that at least some experiences involving music would, either at the time or in retrospect, become autobiographically salient. Those of us who make regular use of media such as film and television are effectively training ourselves to become even better at associating music and life events than we might otherwise be: these media exploit music's ability to trigger memory explicitly to help underscore important moments and highlight aspects of scenes. Research by Boltz, Schulkind, and Kantra (1991) demonstrates systematically what practitioners in the field surely knew, namely that film scenes of which people have no recollection are suddenly remembered again when music is provided as a cue. The effect is stark because a film represents a distinct, definable environment, but the findings of Boltz et al. are equally applicable to the real world; perhaps the efficacy of these media is a self-fulfilling prophesy, as it relies upon and cultivates our propensity to associate music with life events.

Unlike Waterman's (1997) listeners or the questionnaire respondents introduced earlier, for whom associations between music and autobiographical experience were both vivid and explicit, people who associate music with filmic episodes may sometimes find that although aspects of these episodes form an important part of the listening context for future engagements with the same music, the episode itself is not explicitly available. In other words, a listener can forge links between music and experience, thoughts and concepts without even being consciously aware that he or she is doing so; on being revived, these links can be completely decoupled from the details of the original encounter. Nowhere is this phe-

[2]Waterman himself was interesting in attempting to isolate structural features of the musical extracts that might be responsible for emotion evocation.

nomenon demonstrated more clearly than in the advertising world, where soundtracks act as subversive cues, reminding consumers of products and brands (see e.g. Cook, 1998). An amusing illustration is provided by Waterman (1996): one of the participants in his experiment declared that a musical extract being used as a stimulus reminded her of a car speeding through the countryside; it was later discovered that this extract had been used previously as the soundtrack to a car advertisement, where the model in question could be seen gracefully speeding along an open country road. Here, the association between music and image was explicit enough, but the listener had no idea of why she had formed that association.

North, Hargreaves, and McKendrick (1999) present an example of implicit associations between music and a product where not only the rationale for formation, but the actual association itself remains quite outside of conscious awareness: over an extended period, they observed the effects of French and German music on sales of French and German wine in a supermarket; they found that playing the music of the respective countries had a marked positive effect. Of particular interest here is the fact that consumers who were questioned after their purchases seemed blissfully unaware of the way in which their wine choice had been manipulated.[3] Presumably, the music provided a cognitive context in which matching wine seemed somehow suitable for purchase. Although North et al. do not comment on this themselves, it is probably safe to assume that the converse is also true, namely, that the act of wine purchasing provided a context within which the music itself was heard.

There are many reasons why different people might respond to different music in different ways (see e.g. Kemp, 1997) and, as Sperber and Wilson (1995, p.16) observed, "differences in life history necessarily lead to differences in memorised information .... While grammars neutralise the differences between dissimilar experiences, cognition and memory superimpose differences even on common experiences". The effects of autobiographical memories, as documented by Baumgartner (1992), Waterman (1997) and Gabrielsson and Lindström (2000) stand testament to the ramifications of these differences for a listening experience. For most people, however, a life history does not consist solely of a string of isolated autobiographical memories, but also involves immersion in a culture; the findings of North et al. (1999) and Waterman (1996) provide examples of links between music and the extra-musical world that are not so much personal, but culturally shared. According to Gavin and Mandler

[3]Similar effects have been found by many other researchers, such as Yalch (1993), who have observed that shopping behaviour is frequently influenced by store music without shoppers being aware that this is so.

(1987), "to listen to a piece of music is to be engaged in a constant process of interpreting it through the activation of relevant schemas"(p.267); although Gavin and Mandler were primarily concerned with schemata required to understand the music of a certain idiom, their observation could apply equally well to schemata that constitute a listening context seeded from culturally shared associations.

Culturally shared associations need not necessarily be those insidious, implicit ones of whose sources listeners are not aware: declarative knowledge about the circumstances surrounding the inception of a piece of music, or awareness of its significance to other individuals or societies are examples of culturally shared, explicitly available information that could profoundly affect the context in which a listener hears a piece of music, and thereby the range of responses that the piece might evoke. For example, Strauss's *Metamorphosen* is a romantic work for twenty-three solo stringed instruments; it is not explicitly "programme music" and is not ostensibly "about" anything. Despite this, knowledge that the work was an elegy written by an elderly man witnessing the destruction of his culture and his country by the Nazis (see e.g. Kennedy, 1995) may well provoke a listener into mediating their response to the music with considerations of war and suffering. Similar examples abound in the popular music repertoire, where knowledge of the circumstances of music's inception or adoption can be utterly crucial to a listening experience. In terms of the Sperber and Wilson (1995) description, such knowledge, just like autobiographical associations, constitutes relevant assumptions about the world within which the music is heard.

All the observations made so far can be seen as examples of *indexical association* as defined by Dowling and Harwood (1986, see Chapter One of this thesis); research that has been cited, then, can be interpreted as evidence that indexical associations do indeed, as Dowling and Harwood suggest, form an important part of a listening experience. A significant feature of these associations is that the music itself is always irrelevant beyond its service as a trigger for a memory or, in the case of declarative knowledge about a work's significance, a context; in most cases, had a totally different piece of music been involved in the original experience, or perhaps had Strauss had written a different work in the same circumstances that he wrote *Metamorphosen*, the contextual associations would be no different. There are several instances of musically sparked association, though, which display a rather more complex relationship between music and its associand. For example, every country, every culture, and even every generation has its music; national anthems, football tunes, well-known hymns and brass band marches can have a specific significance to a specific group of people, and sometimes, an affinity with a type of music and an understanding of what it represents can even be the defining feature of a social group (Russell, 1997). Here, the association between music and an event, a mode of thought or a concept is far from coincidental; instead, the specific piece of music is adopted—or composed—quite deliberately for its sound and style. Western national anthems, for example, tend to employ steady, slow rhythms, and are replete with solemn, powerful tunes; that is, they tend to make deliberate use of the arousal mediation properties of sounds and patterns (see Chapter Two) and also to draw for their effect on certain stylistic clichés. These cases, despite relying heavily on context invocation for their effect, do not fit so comfortably under the heading of *indexical association*. Inasmuch as listening to an anthem involves the triggering of associations directly between the music as a holistic concept and its extra-musical significance, the phenomenon is no different from, say, a salient autobiographical memory. Inasmuch as it involves the invocation of the potentially emotive power of musical structures and associations that they may evoke in an accultured listener, however, it seems far closer to the phenomenon described as *Symbol* by Dowling and Harwood; these structural units, after all, only have meaning in a musical sense as part of a complex symbol structure. Yet, to claim that the symbols do not act indexically is an absurdity: it is the very fact that certain structural units in Western tonal music have become *associated* with concepts, moods and attitudes that leads to their employment in anthem music in the first place!

Such a suggestion, of course, is right at the heart of Cooke's (1959) theory of musical language.[4] Just as a television advertisement can provoke the formation of associations in the mind of a listener who may be oblivious both to the formation of the associations in the first place and the invocation of those associations in a new context, repeated exposure to the structures of western tonal music provokes the formation of associations between motifs, idioms and sounds on the one hand, and contexts, moods and situations on the other; these associations can then be invoked automatically and subconsciously to form a context—in the Sperber and Wilson (1995) sense—for listening to music that conforms to those idioms; it is simply a question of the functioning of human memory (see e.g. Anderson and Bower,

[4]Cooke's theory is not discussed in full in this thesis, although elements of it have been mentioned previously (see Chapter Three). For a more thorough discussion of the theory, see Agawu (1999).

1973). Seeking a rationale for the age-old question of why minor keys are sad, Robertson (1934) suggested that the process of musical acculturation itself can be attributed to the cumulative effect of association formation. In a paper that foreshadows Cooke, she claims that composers' perpetual use of minor keys in music with programmatically appropriate titles was "bound to give rise to a convention which ... eventually applied to all music"(p.205).

Further evidence of the tokenistic, associationist role of conventional musical idioms in music perception is provided by recent research by Krumhansl (1998), who investigated listeners' ability to recognise and respond to conventional *topics* in classical music. Specifically, she chose two pieces of string chamber music by Mozart and Beethoven whose *topics* had been previously analysed by Agawu (1991), and asked participants in her study to rate the pieces in real time on a number of dimensions. Results showed that judgements of listeners, who were not all musically trained, correlated well with the topics that Agawu had identified in the pieces, leading to the conclusion that these topics "have psychological reality"(p.132); in addition, they were implicated in listeners' emotion judgements. Almost two decades before Krumhansl's study, Maher (1980) carried out a similar, but much more limited investigation of musical intervals and listeners' emotional responses to them. Whilst this study lacked a real musical context and, by current standards, would not be considered even remotely ecologically valid, it did show that some intervals consistently evoked emotional reactions in listeners, and that those reactions "were generally found to be consistent with popular notions concerning the effects of various musical intervals"(p.320). In other words, whether they are aware of it or not, listeners do recognise and respond to conventional units in music. Even response to the musical structure of a work can be understood, in part, simply as response to a listening context. In the light of these observations, the implicitly formed associations between music and filmic episodes or television advertisements discussed earlier can now be seen as exemplifying relationships between musical cues and extra-musical concepts on the level of motifs and patterns; it is not the music as a whole to which the associations are attached, but those musical symbols that characterise a style or genre.

Most of this chapter so far has been concerned with facets of a listening context that derive from associations between music and concepts that lend it emotional salience. Before ending this section, however, it is worth noting that there are facets of context that can impinge on the emotionality of a musical experience that derive not from the music itself, but from the listening situation. Just as knowledge of the circumstances surrounding the inception of a musical work can provide contextual material crucial to a listening experience, so too can the situation in which music is heard. A concert hall, for example, brings with it a set of values, assumptions and expectations that make the experience of listening to a concert very different from listening to a record in a private room! Research by Finnäs (1989) even shows that listeners' expressed musical preferences are very different in public from in private. Modern rock concerts are a well-documented phenomena in themselves, within which mass hysteria and solidarity combine with sense of occasion on a scale so vast that the responses of the average audience member—perhaps "participant" would be a better word—are likely to be largely pre-determined before even a note of music sounds.[5]

In short, an abundance of evidence exists to suggest that listeners associate pieces of music with autobiographical experiences and events and that, once associated, music can act as a cue to trigger the relevant memory. Associations also derive from contextual knowledge of the circumstances of composition of a piece of music, or of the significance attached to it by other individuals and groups. Finally, associations do not necessarily relate exclusively to pieces of music as holistic units, but can also form on the level of musical fragments, so that idioms and conventions themselves create a context within which music is heard and understood. If we follow the lead offered by Sperber and Wilson (1995), then there is no reason to treat any of these as isolated phenomena; instead, all can be understood as the importance of context for hearing, interpreting and understanding music. After all, autobiographical or cultural memory, knowledge of a work's inception or response to a symbol structure all constitute cognitions that form part of the listening context, just as much as does the listening situation itself. The implication of this claim is that if we can account successfully for the role of context in the evocation and mediation of emotional response to music, then we can also account successfully for a vast proportion of the phenomena we might refer to as "emotional response to music". Providing such an account will be purpose of the next section of this chapter.

[5]Lewis (1994) provides a harrowing analysis of two examples of rock concerts where crowd behaviour was so extreme that it ended in tragedy.

Context and the Evocation of Emotion

No hearing person living in the modern world would need to read the evidence cited above to be convinced that listening to music sparks associations, triggers memories and mediates thought. Equally, it hardly requires a thesis to communicate the notion that such associations, memories, and thoughts constitute a listening context that might impinge on emotional response. That music in context—or perhaps music *as* context—can evoke emotion is a truism known to all; the remaining question concerns how it can do so. Chapter Two of this thesis cited evidence to suggest that the sounds of music can cause physiological arousal simply by virtue of neuroanatomical links between the auditory thalamus and the amygdala; for LeDoux (1998) these links constitute a hard-wired survival system, preparing the body for fight or flight in a fraction of the time that it would take the conscious mind to assess a situation and decide upon a course of action. Neuroanatomical links also exist, however, between the amygdala and areas of the cortex (see e.g. Barker, 1991); on a neurophysiological level, then, highly cognitive activities can also lead to physiological arousal which, in conjunction with the cognitions themselves, can constitute an emotional response.[6] Fortunately, elucidating a mechanism whereby associations, memories and thoughts—the concepts of cognition—can trigger emotional responses is not within the remit of a music psychologist; even more fortunately, a formidable literature in the field of emotion addresses exactly this issue.

Of the various explanations for cognitively-evoked emotion that exist, perhaps the most generic and generally accepted is so-called *appraisal* theory, whose argument, broadly, goes as follows: we live in complex groups and societies where circumstances are constantly changing, and where every change carries the potential to have serious repercussions for our well-being. If emotions were all hard-wired adaptations we would constantly find ourselves in situations in which we lacked the mental machinery to survive; given the premise that emotions are functional adaptations, some more flexible mechanism was, in evolutionary terms, inevitable; after all, as Ellsworth (1994, p.151) notes, "in humans, the potential for

[6] There is some debate in the literature over whether emotions require cognitions, or whether arousal alone constitutes an emotional reaction (see e.g. Clore, 1994). A general consensus, however, is that arousal alone is not sufficient and that emotions must consist both of arousal and cognitive explanation for that arousal. The empirical validity of this view was demonstrated in the ubiquitously cited experiment by Schachter and Singer (1962), and is the stance taken in this thesis. The ramifications of this view for the hearing of music as sound and utterance will be discussed later in this chapter.

complexity is as universal as anything". For the appraisal theorist, the problem is solved by an emotion-evoking mechanism at the centre of which is a cognitive system that attempts to determine for everything it encounters the likelihood of that thing impinging either positively or negatively on the well-being of the person who encounters it. The outcome of the appraisal process, and the certainty with which a conclusion can be drawn, determines the nature of the emotion evoked. Perhaps the most sophisticated formalisation of appraisal theory is that offered by Ortony, Clore, and Collins (1988), for whom everything in the world is considered either as an *event*, an *agent* or an *object*. Emotion is always a valenced reaction to one of these categories: one experiences pleasure or displeasure at the outcome—predicted or actual—of an event, approval or disapproval of an agent, and liking or disliking of an object; these characterisations derive from the desirability of an event in relation to one's goals, the praiseworthiness of an agent in relation to one's standards, or the appealingness of an object with respect to one's attitudes or tastes. Ortony et al. use these simple concepts to build a sophisticated model that can account convincingly for the way in which highly cognitive activity could lead to the evocation of emotions.

Affective reactions to the associations, memories and thoughts sparked by music do not always fit comfortably within the scope of "emotion" as described by appraisal theory for two reasons: first, emotions are assumed primarily to be immediate and instantaneous responses to a situation because they prepare action-readiness (e.g. Frijda, 1986), something which musically-sparked cognitions are unlikely to do; second, emotions are assumed to be short-lived and have clearly specifiable antecedents (e.g. Ekman, 1994a), whereas musically-sparked cognitions can be vague and last for the duration of a listening experience, or perhaps even beyond. In terms of emotion theory, the type of affective experience normally associated with musically-sparked cognitions are instances of a related phenomenon normally referred to as *moods*, which are longer lasting, not necessarily linked to specific antecedents, and can be "produced in cumulative fashion over time"(Davidson, 1994, p.53).[7] In many appraisal theories, moods serve as the affective background against which new stimuli are understood; in other words, moods provide a context for cognition.

[7]It is worth noting that the distinction between emotion and mood, although intuitively satisfactory, is somewhat informal; theorists make distinctions based on differing criteria, and some even choose to make no distinction. As Lazarus (1994) noted, in the considerable literature on the relationship between affective phenomena and memory the distinction has barely been made at all: "I will use the terms mood and emotion interchangeably because they have never been distinguished in the research performed on this problem"(p.306).

Chapter Two of this thesis reviews evidence to suggest that repetitive patterns found in some music can invoke mood directly by affecting physiological functioning; in addition, we might expect that a listening experience should be able to mediate mood at a highly cognitive level, regardless of acoustics, simply by virtue of its context of associations, memories and thoughts. A wealth of anecdotal evidence and empirical study suggests that this is indeed the case. That autobiographical memories can affect mood is a well documented phenomenon (see e.g. Singer, 1970; Pillemer, 1993);[8] both in the laboratory and in therapy, mood induction procedures such as the Velten test rely for their efficacy on the ability of thoughts, memories and associations to reliably induce mood (Velten, 1968; see also Finegan and Seligman, 1995). As for music, Albersnagel (1988) has shown that musical stimuli can induce mood more successfully than Velten tests, and studies by researchers such as Stratton and Zalanowski (1989, 1991) suggest that its power as a mood induction tool derives from context-related cognition. Extensive free reports from listeners, such as those collected by Gabrielsson and Lindström (2000) also illustrate vividly how a context created by the experience of listening to music has the power to induce mood.

Whereas emotions bias action, moods affect cognition (Davidson, 1994); further, mood affects memory recall (Bower, 1981) and cognitive flexibility (Isen, 1987). The implication of these observations is that mood itself constitutes an important part of the context within music is heard and, therefore, has the potential profoundly to affect a listening experience. A graphic illustration of this phenomenon is provided by Konečni (1975) and Konečni, Crozier, and Doob (1976)—both reviewed in Konečni (1982)—in a set of experiments designed to investigate the relationship between mood and aesthetic choice. Using a variety of procedures ranging from interjecting musical extracts with loud, hideous noises to being objectionably rude, the researchers induced anger and irritation in participants in the study; they observed that listeners who had been subjected to this barrage displayed strong preferences for simple, non-challenging melodies over more complex ones. A different but complementary phenomenon has been noted by Holbrook and Schindler (1989) who, in their investigation

[8]That memory can affect mood is an issue that has been somewhat perplexing for appraisal theorists, as memories—particularly distant ones—do not obviously fall into the category of issues having immediate relevance for survival! Lazarus (1994, p.308) suggests that this apparent vagary can be understood because "memory does not occur in a vacuum, but is generated in an adaptational encounter or setting in which it is functionally connected with what is now happening. Most commonly ... one is struggling with one of the classic human problems ... which are apt to be recurrent. Most such problems in adults are not totally new, except in certain details, but reinstate struggles of the past that have not been resolved".

of popular music preference, found that preference for particular works was often associated with autobiographical memory, particularly of salient emotional memories of rites of passage.[9] Konečni (1982, p.500) cites such findings as illustrations of the need "to analyse preference with at least some reference to the typical situations in which music is appreciated ... in the stream of daily life"; a suggestion compatible with Konečni might be that music must be understood in terms of the context—conceptual and affective—in which it is heard.

An overwhelming amount of emotion research, then, both theoretical and practical, supports the widely held intuition that musically sparked associations, memories and thoughts mediate mood and, by extension, have a significant effect upon the affective tenor of a listening experience; no holistic model of emotional response to music, then, can afford to ignore the effects of these contextual phenomena. It was claimed in the first half of this chapter, and previously in Chapter One, that elements of musical structure—cadences, motifs, and common patterns—are effectively tokens that can also spark associations, memories and thoughts that become part of a listening context in just the same way as can whole musical works. If this is indeed the case, we should expect that the presence of these elements in music would have an important part to play in the evocation of emotion. In fact, these elements have probably played a crucial role in work already cited: if Albersnagel's (1988) listeners did not actually know the music that he played them during mood induction, the effectiveness of the procedure must have been, at least in part, a result of emotional response to the *type* of music, that is, response to a recognised set of musical symbols which—presumably by process of acculturation—had become *associated* for those listeners with emotional valence; likewise, it was surely on an atomistic level that Waterman's (1996) listener associated a musical extract with a car breezing down a country road. Certainly, such an interpretation is congruent with Krumhansl's (1998, cited earlier) findings. In an extraordinary experiment whose aim was to study the effects of music on visual imagery, Osborne (1981) asked participants to lie comfortably in a room and listen to music that consisted of synthesised, "spacy" music of unknown origin; once the music was over, all the listeners reported freeform on their experience. Results showed all reports to be remarkably similar, involving

[9] Other anthopologically-oriented researchers have noted that listening context can affect aesthetic judgements independently of mood. For example, Boris (1993, pp. 271-272) emphasises the importance of cultural value on musical preference when he states that "since it is in the nature of cultures to be discriminate, our sensitivity imbibes values and tastes as well as techniques and meaning".

descriptions of "natural scenes", "religious images" and "outer body experiences". A similar experiment, carried out by Quittner and Glueckhauf (1983), found a very similar patterns of results. Whilst part of this response was probably predetermined by the rather peculiar experimental situation which for many listeners, no doubt, sparked associations—or perhaps imaginings!—of a hippie-commune or some similar concept, surely it was also due, in part, to the music itself: that is, its sounds, its structures and the associations sparked by these *intra-musical* elements.

Perhaps it is not possible to determine for sure what the comparative contribution of listening situation and the music itself might be in terms of sparking associations that may be responsible for mood evocation; certainly, there appears to be no empirical work that sheds light on the issue. According to the model of emotional response to music being presented in this thesis, however, the answer to this particular question is of little significance. To distinguish between associations sparked by a listening context and those sparked by musical symbols in terms of evaluating contribution to emotional response is no more useful than to distinguish between, say, knowledge of the circumstances surrounding the inception of a work and awareness of that work's significance to a group or culture; all of these things simply become part of the context in which the music is heard and, as such, have the power to mediate mood.

Even though there is no meaningful difference between associations sparked by a musical token and one sparked by a piece as a whole in terms of emotional response to music in terms of their contribution to listening context, there certainly is a difference between the objects in themselves: whereas a whole piece of music typically lasts for several minutes, a single symbol may sound for a mere second. A question arises, then, as to what the combined effect of several symbols and their associations may be. Two possibilities present themselves: first, that the effect of symbols is cumulative, so it is not just a single token but a sequence in combination that is likely to spark associations that become part of the listening context;[10] second, that a single, isolated token forms part of a context only fleetingly, until supplanted by the next. In other words, we could hypothesize a two-tiered system of context involving static, global components and dynamic, local ones. The static, global

[10]It could, of course, be argued that a combination of symbols is in itself just a single larger-scale symbol. Such an argument would be congruent with hierarchical conceptions of musical structure perception (e.g. Lerdahl and Jackendoff, 1983).

context, as we have already seen, can lead to mood induction, but it is not so clear that dynamic local ones would have the same effect: inasmuch as the local lead to a cumulative static global amalgam, we can hypothesize that their role is identical to that of any other global aspect of context, but inasmuch as they are fleeting ephemera, they are much more likely to lead to transient *emotions*. Davidson (1994, p.52) suggests that "moods provide the affective background, the emotional color, to all that we do. Emotions can be viewed as phasic perturbations that are superimposed on this background activity"; likewise, static associations sparked by music and a listening situation are a context, a cognitive background, whereas ephemeral symbols and the moment-by-moment dynamics of their interactions are perturbations superimposed on this context. Ramifications of this conceptualisation will be discussed in the next chapter.

The idea that a set of transient symbols may contribute to listening context through some sort of cumulative effect suggests a way in which sounds and utterances too can contribute. We have seen in the previous two chapters how listening to music as sound and utterance can evoke emotion by virtue of the acoustic properties of those sounds and utterances and the way in which our perceptual and emotion systems work. In addition to these effects, it seems likely that sounds and utterances could contribute to the experience of listening at a cognitive level: we know what the sound of a gun is and what its significance might be; we can recognise an anguished cry and can relate it to conceptualisations of human suffering. It is not hard to imagine how use of these or similar sounds and utterances in music might spark a host of associations that crucially change the cognitive context within which the music is understood, thereby dramatically affecting emotional response to that music.

Sound, utterance and context may interact in another way too, as can be seen by invoking Schachter and Singer (1962) once more, this time in more depth. In this experiment, some participants were injected with a serum—adrenaline—and warned that they may feel physiological arousal as a result; other participants were also injected but were tricked into believing that the injection would have no such physiological effect. Both sets of participants were then put in social situations designed either to amuse or infuriate, with extraordinary results: those participants who had been warned of the possible effects of the injection reported being unmoved by the social situation; by contrast, those who had not been warned reported being either highly amused or extremely angry, depending on the situation in which they were placed. This brief description does little justice to what was a highly sophisticated experiment, but the crucial point can nonetheless be seen: when participants could not attribute a feeling of physiological arousal to experimental intervention, they rationalised it to themselves in terms of context-evoked emotion. In other words, context determines the emotional effect of physiological response. We have seen in this thesis that loud sounds or those with certain timbral characteristics can cause physiological arousal independently of cognition; in the light of the Schachter and Singer experiment, we can also suggest that the context in which music is heard could provide the cognitive component that turns this physiological arousal into emotion.[11]

Given the relative dearth of research that has explicitly examined musical structure in terms of associations and the way in which various facets of listening context affect the emotional tenor of a listening experience, some of the suggestions presented here concerning relationships between these domains must remain somewhat speculative. It should be utterly clear, however, as the epigraph to this chapter asserts, that music is always heard in context. During the listening process, associations, thoughts and memories deriving from the music itself, from knowledge about its significance and from contingencies of the situation in which it is heard form a rich and changing context that is central to emotional response. According to Kivy (1989, p.157):

> "Of course, the *Grosse Fuge* might arouse anger in some listener or other by reminding him of an angry encounter with his boss; so too, however, might the tranquil strains of the *Spring Sonata* .... This is a familiar enough phenomenon; but it has nothing whatsoever to do with the expressiveness of the music except by accident".

In many ways, this very well may be true: listening context is extraneous to the musical object itself, even if some other elements, such as the context in which a piece is written, may not be. The extent to which context may or may not be relevant to an understanding of the aesthetics of art music, however, is of no concern to us here. An overwhelming volume of evidence indicates that context is crucial to a listening experience and to emotions evoked by such an experience; context is therefore something that no model of emotional response

[11] The interaction between aspects of sound, utterance, local musical structure and context will be considered in detail in the next chapter.

to music can ignore. As Konečni (1982, p.501) shrewdly observed, listeners are "engaged in a constant exchange with the social and non-social environment, of which the acoustic stimuli are a part". Hence, that music is heard in context is the third premise of the model presented in this thesis.

Chapter 5

Music as Narrative

> *"I think I have experienced your symphony. I felt the struggle for illusions; I felt the pain of one disillusioned; I saw the forces of evil and good contending; I saw a man in a torment of emotion exerting himself to gain inner harmony. I sensed a human being, a drama,* truth, *the most ruthless truth"*(Mahler,1904).[1]

The above quotation is an extract from a letter written by Schönberg after hearing the first performance of Mahler's third symphony. In many ways, the poetic rhetoric and a tone that verges on the melodramatic are symptomatic of the cultural climate in which the letter—and the symphony—was written (see e.g. Schorske, 1979). Schönberg's response to the symphony, however, reveals something rather fundamental about a mode of engagement with music whose significance extends way beyond the immediate cultural context of that letter: embedded within his references to truth and struggles, to disillusionment, good and evil is the implication that music can represent or evoke the representation of such concepts. Embodied within the symphony, apparently, are "protagonists" who can undergo trials and tribulations, who can struggle, who are human. In short, the symphony is not so much a piece of music, but an entire narrative world.

Intellectually, the idea that a symphony is a narrative—or at least that it can be heard as one—seems rather bizarre; it is certainly remarkable enough to have kindled endless philosophical debates (see e.g. Treitler, 1993). Despite this, Schönberg is hardly unique in described music as though it were indeed a kind of narrative: the music theoretic literature

[1]This excerpt is taken from a letter published in a collection of letters and anecdotes compiled by Alma, Mahler's wife (Mahler, 1968, pp.256-7).

is full of narrative metaphors and allusions, and has been so for centuries. For example, as early as the ninth century, the theorist Boethius made the following statement about dissonance in his treatise, *De institutione musica*: "as long as they [dissonances] are unwilling to blend together and each somehow strives to be heard unimpaired, each is transmitted to the senses unpleasantly" (Boethius, 1989, p.16). Dissonant notes here are imbued with a "life" of their own, and play a specific part in a narrative; they are deliberate antagonisers of the musical calm. Such metaphors have persisted even until the present day: at this end of the temporal spectrum, theorists such as McClary (1991) have made extensive use of similar narrative and personification metaphors in their analyses of musical forms.

It will come as no surprise to most people who listen to music that musicians and music theorists have consistently turned to the concept of narrative in their discussions, because intellectual problematisation of the issue notwithstanding, it is virtually a folk-psychological truism that music has narrative qualities. One very good reason for this is that much of the music to which we are exposed has an *explicit* narrative component: folksongs, film soundtracks, programme music, opera and oratorio are all examples of genres within which music and narrative are inextricably linked. Given the quantity of music of this sort to which people are typically exposed, it is perhaps unsurprising that the same modes of description—and possibly modes of thinking—extend to music that does not have this explicit narrative element. Another plausible reason is that narrative modes of explanation reflect something fundamental about the way in which we make sense of the world; that our responses to music and the way we understand them are mediated by the "narrative structure of experience" (Shweder, 1994, p.42). In the light of these observations, rather more bizarre than the notion that music might be considered as narrative in some sense is the fact that music psychologists studying emotion have refrained almost entirely from thinking of music in this way, or from drawing on the extensive literature concerning narrative in other domains.

This chapter is concerned with the relationship between music and narrative, and its implications for emotional response to music. It will start by discussing various ways in which music can be said to possess narrative content, and some of the difficulties facing theorists who have attempted to explicate the nature of musical narratives, particularly in western art music. It will then move away from the largely musicological and philosophical arguments to demonstrate that from a psychological perspective, conceptualisation of a listening experience as narrative comprehension provides a powerful means to understand two highly salient attributes of emotional response to music, namely dynamism and coherence. The dynamic nature of emotional response to music is well known: arousal waxes and wanes through moments of intensity and moments of calm; it is no accident that Meyer's (1956) account of emotion as response to a series of tensions and resolutions is about the least controversial and the most generally accepted theory of emotional response to western music to date. Much of the second half of this chapter will be devoted to a consideration of these dynamics as narrative process. It will attempt to show that far from being a phenomenon peculiar to music, these dynamics can be understood as a specific instantiation of a far more general phenomenon. Emotional response to these dynamics will be seen to be the result not so much of inhibition of tendencies (see Chapter One) but as a natural outcome of narrative process. In addition to being highly dynamic, emotional response to music also appears to be remarkably coherent: published reports of listener reactions do not talk of listeners experiencing series of disconnected emotions in response to disjunct fragments; instead it seems that however complex a listener's responses to a piece of music in a given context, disparate fragments are somehow bound into a unified experience. The final section of this chapter will be devoted to consideration of narrative comprehension as a mechanism for binding response to sound, to utterance and context into a dynamic coherent experience.

In contrast to the previous chapters, the discussion of narrative here will mention very little empirical work from within the field of music psychology. The reason for this is that hardly any empirical work has addressed the possibility that emotional response to music might be emotional response to narrative, or even that music may be heard as narrative. This chapter—and the remainder of the thesis—will argue that rectifying this situation could help substantially to increase our understanding of emotional response to music.

Narratives in Music?

A problem that perpetually plagues discussion of the relationship between narrative and music is one of defining the phenomenon under study; as Noy (1993, p.13) asks, "if music can be interpreted as a narrative, what is the nature of the musical narrative?". Part of the problem lies in the word "narrative" itself, whose definition is rather elusive even before any attempt is made to apply it to music; mostly, however, the difficulty concerns the application of the idea of narrative in the musical domain. For theorists such as Labov (1972), a narrative is any sequence of clauses that match a sequence of events; thus any two clauses that follow on from one another in temporal succession can be deemed to constitute a narrative. There is nothing problematic or ambiguous about that definition, which has a clear, tight scope. Unfortunately, however, it is of little use to anyone but a linguist, and leaves no room for a phenomenon such as narrative in music; neither does it encompass much of what most people would understand narrative to be. By contrast, Bruner (1986, p.13) provides the following description of what he calls the *narrative mode*:

> "It deals in human or human-like intention and action and the vicissitudes and consequences that mark their course. It strives to put its timeless miracles into the particulars of experience, and to locate the experience in time and place".

Whilst this description hardly constitutes a formal definition of narrative, it does serve to illustrate the wide scope of applications that the word has enjoyed; it also reflects more closely than definitions such as Labov's what a "folk-psychological" view of narrative might be.

Slightly less poetic, but equally broad in scope is Gerrig's (1993) description of *narrative worlds*, places to where a reader gets "transported" by a story, or even just by a concept or idea. For Gerrig, the mental process of transportation is far more interesting to study than the specificities of the texts which cause that process, a proposition to which we will turn later on.

There are many musical forms in which words play a prominent role by virtue of which they have a clear narrative structure. Songs, operas and musicals are often literally narratives even in the strict Labovian sense of the word: they consist of a sequence of clauses that refer to events, situations or thoughts. Even so, the relationship between the words and the music—and therefore the nature of the narrative—can vary considerably from style to style and from work to work. At one end of the spectrum, for example, might be a relationship in which the entire narrative content is focused on the text, music playing a merely accompanimental role. In this category could be a song in which music is constructed without any regard for the text, perhaps by mechanistically repeating a simple, well-worn chord progression or rhythmic pattern. In such a situation, music and narrative would merely exist contingently, and in no sense would there be a "musical" narrative to be perceived by a listener or debated by a theorist. Of course, in reality a relationship such as this is extremely rare, because even the most mundane song is likely to present a more sophisticated relationship between narrative and music. If any musical concession at all is made toward the text, or any attempt is made by the music to "reflect" the textual narrative in any way, it can no longer be said that that the two domains merely co-exist; instead, they become inextricably related. In fact, from a listener's perspective, even if nuances in the music are only *perceived* to be reflecting the narrative of a text, as is inevitable if the music and text are perceived as a coherent whole, the relationship becomes more complex.[2]

Even in more realistic cases, the fundamental relationship between a text and music may still remain relatively straightforward. For example, the music of many pop songs is written according to a limited number of well-defined idioms: certain types of harmony, instrumentation and rhythmic pattern are chosen depending on the general mood of a text, and melodic contours and occasional changes in instrumental texture will follow textual specificities, again according to loosely prescribed formulae (see e.g. Citron, 1985). The relationship between text and music is somewhat asymmetric here in that the textual narrative dictates the musical form; however, once written, the music alone is able to convey at least some of the narrative content of the text—particularly the emotional content—to an acculturated listener.[3]

Monteverdi's final opera, *L'incoronazione di Poppea*, is an example of a work in the western art music repertory that presents a rather different balance of influence: here, the music does not merely respond closely to the text, but also constructs a narrative of its own by providing a commentary on the textual narrative (Fenlon and Miller, 1992). In the first scene, fragmentary vocal lines and circular harmonies mimic the seductive Poppea who urges Nero

[2]In a similar vein, Cohen (1999) has suggested that one of the primary roles of music in early film was to mask extraneous noise; for example, it was used for "drowning out the sound of the film projector"(p.53). She too notes that even when fulfilling such a mundane function, the music is heard as part of the film and was always chosen to be congruent with the mood of the filmic narrative.

[3]Such a suggestion is congruent with the notion of musical tokens sparking associations (as presented in Chapter Four). But see also Stratton and Zalanowski (1994), whose empirical study indicates that perceived emotional expression of certain pop songs varies depending on the text to which the music is set. This finding is reminiscent of Hanslick (1891) in his notorious commentory on the music of Gluck's aria, *Che Farò senza Euridice*, where he claims that the music only sounds so tragically sad as a result of the words, and that a different text could make it seem profoundly happy. Sadly, both the aria itself and Hanslick's comment are so well known that an empirical test of this hypothesis could never be effectively executed!

not to leave her. This merely provides a paradigmatic example of music supporting a text within the constraints of a particular musical style: the notes "shadow" the narrative as it progresses, mimicking in the musical domain what the words are doing in the textual. A few scenes later on, however, the relationship between these two domains changes radically: Seneca sings of the importance of steadfastness, declaming constantly on a single pitch; meanwhile, the instrumental bass-line descends, moving away below his metaphorical feet! Here, the music actually undermines Seneca and by so doing totally changes our perception of the scene: Seneca is telling us one thing, but the music is telling us something quite different. In neither of the above scenes can it be said that the music narrates independently from the text—it would be a powerless narrator if the text were removed—but given that the text is present, the composer's deliberate anti-conventional application of musical features makes an independent statement of its own. *L'incoronazione di Poppea* presents a story that would not be complete without the music; it must contain, therefore, a narrative in music. To be sure, understanding the musical narrative embedded within Monteverdi's opera requires a high degree of familiarity with the idioms of the musical style—a listener without an inkling of the conventional musical treatment would not notice the subversion of Seneca—but in this regard, music is not dissimilar from language, where a listener requires quite a considerable amount of knowledge before he is able to understand narratives communicated in that language.[4]

So far, discussion has focused only on musical genres in which words and music sound concurrently. There are plenty of common musical genres, particularly in the western art music tradition, however, where music sounds without text but still contains a strong narrative element. An obvious example is programme music, purely instrumental that which openly declares its intention to represent a scene or a tale by virtue of the "programme", a textual narrative on which the work is based. The exact relationship between music and text depends on the specific work and the composer's intentions, but to the extent that music can represent a text at all, it will surely do so via imitation and idiomatic conventions. Just as

[4]This consideration of an accultured listener's sensitivity to musical idioms is not a far cry from an equivalent discussion in Chapter Four, where it was established—as Cooke (1959) suggested—that accultured listeners can indeed recognise and respond to conventional idioms in the music of their culture. It was further suggested in that chapter that musical "tokens", once recognised, can spark associations by virtue of their normative usage, which then become part of the listening context. Here, it can be seen that in a specific narrative context such as Seneca's monologue, these tokens and the musical situations with which they are normally associated can actually carry narrative meaning, or at least contribute to a meaningful narrative complex.

the music in *Poppea* only makes narrative sense when the text is present, it is reasonable to suppose that the musical narrative of a programme symphony requires its programme in order to make full sense of the narrative structure. Unfortunately, musicological debate about programme music appears to be occupied almost exclusively with the question of whether or not such music can ever make sense in the absence of its programme and, if so, what the implications might be for such music as art (see e.g. Treitler, 1980, 1984). While this issue—which is of no interest to us here—remains unanswered, questions concerning the nature of a relationship between music and narrative appear to attract scant attention.[5] Luckily, there is a genre closely related to programme music where this debate does not exist at all, namely film music. As film music does not hold the same cultural status as western art music, it tends not to be subjected to the same degree of ontological scrutiny: nobody worries about whether it can stand as music apart from its film; to do so would probably be considered rather perverse. This makes it a rather useful medium for consideration of the relationship between music and narrative.

In many ways, a film score is no different from other types of programme music: essentially, it overtly follows a textual narrative and employs a number of conventional devices to do so. On the whole, film music consists of localised fragments rather than large-scale structures, and always reflects specific moments in a plot—perhaps making it more analogous to opera than a symphony in this respect—but the idea is similar enough. Recent research by Bullerjahn and Gueldenring (1994) demonstrates vividly that a musical soundtrack dramatically alters not only the emotional tenor of the film clip it accompanies, but also the relating of the filmic narrative itself. In their experiment, participants watched a ten-minute film that was created to give the impression of being an excerpt from a thriller; typical clichés, such as mysterious cuts between an elderly man wearing a trench coat on a train and a lady in a dressing-gown making tea, interspersed with still shots of an alarm clock, a champagne bottle, or a road sign make the genre relatively unambiguous. Participants watched this film to the accompaniment of one of five soundtracks, all of which were written by professional film composers specifically for the experiment; each soundtrack was intended either to paint the impression of a melodrama or a crime thriller, or to be ambiguous. Once the film was over, participants were asked a number of open-ended questions about its atmosphere and

[5] Although the issue of musical narrativity has been picked up in recent years by some exponents of the "New Musicology". (see e.g. Abbate and Parker, 1989; Kramer, 1984).

content: for example, one question was "What is the reason for the journey of the elderly man?". Content analysis of the results revealed that participants' interpretations of the plot was influenced vastly by the soundtrack, or as Bullerjahn and Gueldenring themselves put it, "each musical soundtrack creates its own particular type of film and plot"(p.112). Particularly remarkable was the high degree of intersubjectivity: participants who were subjected to the same condition were highly consistent in their interpretations. In an experiment involving four-hundred participants in total, this was certainly not a foregone conclusion! Surely these findings, totally congruent with everyday experience, are a testament to music's ability to narrate to an acculturated listener by means of the conventional musical devices that typify a genre. So, in short, music does not need a verbal accompaniment to become an integral part of a narrative; the privilege is not reserved exclusively for songs.

Programme symphonies and film scores are not the end of the story, however, as the most conceptually problematic musical genre to deal with in terms of its relationship to narrative is one that has not yet been considered here, namely so-called absolute music—abstract art with no text and no programme. The question that plagues musicologists considering the relationship between music and narrative in the programme symphony pales into insignificance compared with the difficulty facing both musicologists and philosophers trying to discover "whether abstract music, in the absence of extramusical elements can present a narrative line"(Noy, 1993, p.128), and if so, how it can do so. These are recurring questions that have preoccupied countless theorists, who have framed it not so much an issue of the relationship between music and narrative, but as the crux of an understanding of the nature of music in general and of musical emotion in particular. Most of the arguments from the various perspectives are by now very familiar, having been scrutinised thoroughly several times elsewhere (see e.g. Kivy, 1984); nevertheless, a discussion of narrative and music can hardly ignore the debate, so they will be summarised briefly below.

Most theorists have adopted one of two diametrically opposed stances: the referentialist stance or the essentialist stance (see e.g. Meyer, 1956, or more recently Cook, 1998 for a review). Broadly, referentalists claim that music becomes meaningful by virtue of its ability to refer to concepts outside of the realm of music and that meaning lies squarely within the domain of the referands. For example, music can imitate real world sounds such as the twittering of birds or can present motifs that signify something highly specific, such as the sound of a hunting horn, but the "meaning" here belongs to the birds or the hunting horns; a musical narrative emerges as a result of the cumulative effect of the referands. Such a notion is entirely compatible with our earlier discussion here and in Chapter Four. By contrast, to adopt an essentialist stance is to claim that any meaning, and therefore, any narrative, inheres within the music itself, its structures and its forms. Musical patterns of contours evoke emotion due to their formal properties alone; according to Langer (1953), it is the fact that musical contours map the contours of emotional expression that causes musical forms to evoke emotions.

A parallel argument to that between referentialism and essentialism is that between *appearance emotionalism* and *hypothetical emotionalism* (to borrow Davies's 1997 terminology). Levinson (1997, p.28), an *appearance* emotionalist, declares that "much music offers the appearance of human emotion or of persons outwardly manifesting emotional states"; if music is to carry a narrative line, it must presumably do so through the changing emotional states of these "persons". Whilst this contention seems perfectly reasonable, the difficulty is that it begs the question as to whose emotional states appear to be being manifested. In a representational artwork there are always protagonists—characters with whom we can identify—but there is no equivalent in absolute music; the emotion cannot even be ascribed to the composer, as attempts to do this tend to "encounter well-rehearsed difficulties" (Davies, 1997, p.95). For the exponents of appearance emotionalism, however, this is really not a problem. Their stance is very similar to the essentialist one, namely that the ability to evoke emotion is intrinsic to the music itself, thus rendering it unnecessary to attribute the felt emotions to a protagonist. Kivy (1980) invokes the sadness on the face of a St. Bernard dog to illustrate his position: the dog always looks sad simply because of the downward contours of its face, but this does not mean that the dog actually *is* sad; neither does an observer have to assume that the dog is sad in order to detect the appearance of sadness. On the other hand, the hypothetical emotionalist stance (e.g. Mauss, 1988; Newcomb, 1983) claims that a listener will always invoke a virtual persona who is the perceived subject of the emotions. The justification is that appearance emotionalism is totally impotent when it comes to understanding how emotions without specific behavioural expressions can be evoked; this would include many evoked by music.

Proponents of all of these various arguments and positions have perfectly plausible explanations of what emotion in music might be and where it inheres but given that they are incompatible with each other, not all of the explanations can be right. I believe that there are two factors which have led to this seemingly unresolvable problem. First, much of this literature appears to be striving for a single theory that can encapsulate all that there is to understand about emotion, meaning, narrative and their relationship to music. Thus, issues of representation, narrative, meaning, emotion and structure are all blurred together as a single amorphous, opaque complex; differences of opinion between theorists can often be attributed to an implied but unstated stance towards any of the above. The second factor concerns the corpus of music being studied: whilst the nature of representation in absolute music may be an interesting question for musicological philosophy, such art constitutes only a tiny portion of the world's music; the notion of such music does not even exist in many cultures (Arom, 1991; Magrini, 2000). In many non-western cultures, instrumental music always has a narrative function by virtue of its place in ritual or folklore (see e.g. Feld, 1990); the music has an unwritten programme due to its status in the culture of its origin. In the light of research reviewed in Chapter Four of this thesis, we could add that even purely instrumental, absolute western art music always has a *context* — one might call it a programme—that derives from autobiographical and cultural associations that a listener might make, familiarity with a symbol system and even the circumstances in which music is heard. From this wider perspective, it does not seem unreasonable to suggest that in a very real sense "absolute music" does not exist at all. Or as Cook (1998, p.91) put it in his critique of Kivy's position:

> "Do we *ever* hear music alone,[6] and if we do, can we be justified in regarding this as the paradigm case of music listening? A negative answer to both these questions is suggested by the fact that our musical culture invests a great deal of time and effort in an apparent attempt to ensure that we *don't* hear music alone—to ensure that we know Mahler was neurotic, that is to say, and that Elgar epitomised the values of the British Empire in its Indian summer before the first world war".

[6]"Music alone" is a term used by Kivy to refer to listening to music *qua* music without any contingent contextual baggage.

In short, then, there is no reason to suggest that absolute music should be unable to carry a narrative line; rather, it should be able to do so just as can a programme symphony, a film score or even a song, simply by virtue of the context in which it exists and in which it is heard.

This rejection of the concept of absolute music does not merely preclude the embracing of an essentialist stance with respect to musical emotions and narratives, but also reveals a fundamental problem with the very concept of music and narrative as presented in this discussion so far, namely, the conceptualisation of narrative as an *object*, a "thing" that is either present in music or is not. At first glance, this conceptualisation seems sensible enough; after all, people really do seem to *hear* narratives in music and respond to them emotionally, so they must exist in *some* sense. The problem, though, in the context of the present thesis, is that to discuss narrative as if it were an object—to consider the circumstances in which narrative might be said to be contained within music—would be to fall prey to the seductive temptation criticised in Chapter One: it would be to undermine the status of the most important factor in considering any sources of emotional response to music, namely the listener.

The intuition that music is narrative comes from the fact that when we hear music, we adopt a narrative mode of listening. When Bruner (1986, p.13, cited earlier) wrote of "human or human-like intention and action and the vicissitudes and consequences that mark their course" he was not talking of narrative as object, but narrative as a mode of thinking; likewise, Gerrig's (1993) "narrative worlds" emerge not from a text but from mental process. For the music psychologist studying emotion, the questions that plague philosophers and musicologists can be avoided altogether if narrative is considered not in terms of text and music, but in terms of the cognitive processes involved in creating narrative experiences. The question asked at the start of this section—"narratives in music?"—may be interesting for the student of aesthetics, but for us here, it is utterly meaningless. According to Branigan (1992, p.3), "narrative is a perceptual activity that organises data into a special patterns which represents and explains experience"; given this definition, a narrative can never inhere in music, but music *can* be heard as narrative.

The remainder of this chapter will consider what it might mean to hear music as narrative, and what the ramifications might be for the evocation of emotions. In the light of

Branigan's observation, it will be suggested that narrative is an ideal cognitive tool that we employ to help us "comprehend" a listening experience, and that the dynamic quality of emotional response to music can be attributed, in part, to the dynamic nature of unfolding narratives. Bordwell (1985, p.34), in his discussion of filmic narratives, states that "comprehending a narrative requires assigning some coherence"; here, it will be suggested that the coherent nature of emotional response to music can also be accounted for if music and context are understood as narrative. The fact that music psychologists studying emotion have tended to shy away from consideration of music as narrative is perhaps explained by the furious debates within musicological philosophy; a central claim of this chapter, however, is that seen as process rather than object, the concept of narrative can provide invaluable insight into the fundamental nature of emotional response to music.

Suspense and Expectation

When reading a novel, or watching a film, many facets of narrative experience can elicit an emotional response, but some of the strongest responses are those associated with suspense. Questions such as "will the hero save the world?" or "will the evil monster escape?" leave us clutching the edge of our seats or anxiously fingering the corner of a page, hoping or fearing as appropriate. According to goal-based theories of emotions (e.g. Ortony et al., 1988), our emotional responses are governed by the relationship between the situations that we or others find ourselves in, and the goals that we or others are trying to attain (see Chapter Four). What makes emotions such as hope and fear so powerful is the fact that they are complexes involving not just goals and situations but also uncertainty, something which precludes accurate appraisal and therefore undermines our ability to assess a situation with respect to the world.[7] The hold that a skillful narrator has over a reader or listener is often due to the narrator's ability to manipulate levels of uncertainty.

In his extensive work on the subject, Gerrig (1993, p.79) observes that "the experience of suspense should occur when a reader (1) lacks knowledge about (2) some sufficiently

[7] "Fear" is used here in the context of cognitive theories of emotions, to denote appraisal of an uncertain situation whose potential outcome conflicts with goals. To claim in this context that it is a complex emotion is not to contradict the overwhelming evidence that fear is, in fact, a *basic* emotion in the sense that the primate brain contains phylogenetically ancient systems that control behaviour in fight-or-flight situations (see e.g. LeDoux, 1998). The ambiguity here just serves to illustrate the slippery nature of emotion words! (see also Scherer, 1992, who makes a similar comment concerning the troublesome nature of emotion words).

important target outcome".[8] That lack of knowledge is a prerequisite for suspense is not a controversial point; presumably, a "sufficiently important" outcome would be one which has a significant effect on the possible attainment of goals. Here then, in a nutshell, is a perfectly satisfactory explanation for the phenomenon of suspense. One problem that must be addressed, however, is that a reader or listener is seldom the protagonist of a narrative; why should *any* outcome for a character in a story *ever* be of the remotest significance? (Walton, 1990). Gerrig, in common with Ortony et al. (1988), see this as a problem of empathy: if the reader or listener empathises closely enough with the protagonists or the situations in which they find themselves, the requirements for sufficient importance will be met, and suspense will result. Research by José and Brewer (1984) indicates that empathy may well be an intrinsically human phenomenon, by demonstrating that the extent to which even young children worry about the outcome of stories is directly correlated with the extent to which those stories focus on protagonists with whom they can empathise.[9]

Gerrig (1993, p.79) makes two further claims about suspense in narratives, namely that "feelings of suspense will be heightened to the extent that (3) the target outcome maps out a challenging problem space and (4) the author is able to sustain participatory responses over a period of delay". In one sense, the "problem space" belongs to the characters; it defines the parameters of the situation in which they find themselves and sets the limits on possible outcomes. The question of how challenging it may be, therefore, can be mapped simply onto a question about the uncertainty of a target outcome. It is not only the characters, however, but also readers and listeners who embark on a problem-solving exercise: part of the experience of narrative is to wade through potential options, to consider potential escape routes, or to weigh up moral dilemmas. In a very real sense, the problem-space belongs to the reader; suspense is heightened not merely through empathy, but through anticipating the outcome of one's own calculations! The significance of Gerrig's fourth point, that delay heightens suspense, is merely that the longer a delay before a problem's resolution, the longer uncertainty reigns and, quite simply, the longer suspense lasts.

Music that incorporates textual narrative, such as an opera or song, can be expected to evoke suspense exactly in the way outlined above, provided that the text concerned is a

[8] Gerrig (1993) presents an extremely comprehensive account of narrative experiences and the way they are evoked by reading and other similar activities, such as watching a film. Hence, the text will be cited frequently throughout the rest of this chapter.

[9] Such a notion is also compatible with the idea of "social intelligence" (see e.g. Gallup Jr., 1998).

suspenseful one. For that matter, even if the text is not particularly suspenseful, one might expect empathetic feelings towards a protagonist to lead to curiosity about them that extends beyond the boundaries of the text itself.[10] Of more interest to our present concerns, however, is the fact that there exists a striking similarity between the evocation of suspense in narrative as formulated by Gerrig and a musical phenomenon that need not be connected at all with an external text or story, namely the phenomenon of expectation. Replacing the concepts of fear and hope with that of expectation reveals the possibility that narrative suspense has a direct equivalent in the musical domain. As we have already seen (Chapters One and Two), the act of listening to music involves generation of expectations—assumptions about how the music will unfold over time—that guide the listening process and thus, in turn, mediate affective response. Some of these expectations are generated by the workings of the *gestalt* principles, rules governing perception of patterns; yet others work on a much higher level, being derived from a listener's knowledge of a particular musical style or even an individual work. In terms of Gerrig's analysis, they constitute a structure that has the potential to fulfil the four requirements for the generation of suspense: the primary requirements are lack of knowledge and a sufficiently important outcome, both of which can be seen to be equally important to the generation of expectancies in music. Furthermore, their effect is constrained by similar limitations as those that constrain their equivalents in the textual domain.

Beyond those expectancies that derive directly from *gestalt* principles, any expectancy generated by a listener in response to a piece of music must be dependent on the listener's schematic knowledge of the musical style, just as an author's ability to generate suspense in a narrative depends on the fact "that they share a common background ... with their audiences" (Carroll, 1997, p.205). A listener unfamiliar with a style will not generate many high-level expectancies regarding the continuation of a piece of music within that style, just as a reader who does not know that evil monsters are dangerous is unlikely to be too concerned as to whether or not one escapes from the zoo! In the context of this background knowledge, expectancies are generated (a problem-space is constructed), but as music (or a narrative) unfolds, the listener never knows for sure that the expectancies will be borne out; in other words, expectancies leave plenty of room for uncertainty. An obvious example in western tonal music of an expectancy that is not borne out is the aptly-named interrupted

[10]This is a common effect of immersion in a narrative (Gerrig, 1993), something that will be discussed further in the final section of the chapter.

cadence: background knowledge of the musical idiom would lead a listener to expect a tonic chord at the end of a cadential passage; however, an outcome is never certain until the chord actually sounds, so there is also a possibility, albeit an unlikely one, that another totally different chord will sound. In the event, a different chord—the submediant—does sound; in retrospect, the cadential buildup has led to an expectancy violation, an effect that has been made possible only because of an implied uncertainty. If the listener knew for certain—and therefore expected—that a submediant chord would sound, its occurrence would be thoroughly unremarkable. If listeners *always* knew for certain what the next musical event would be, the listening experience would probably be rather dull, at least from a narrative perspective.[11]

At first glance, the concept of a "sufficiently important target outcome" for suspense generation lends itself rather less well to musical analogies. One possible line of argument could be that a sense of importance is "borrowed" from the textual narrative aspect of music, such as the plot of an opera, the words of a song or the programme in a programme symphony; in other words, the extent to which it matters whether expectancies are confirmed or violated in the musical domain is mediated by the extent to which a listener empathises with the characters or situations of the text. This argument could be extended to purely instrumental music on the basis that music always has some narrative content, even if it derives from a listening situation, and so the listener can always find something with which to empathise. An alternative, possibly complementary hypothesis might be that sufficient importance derives not from the music but from the listening context or the mood that the context has evoked in a listener (see Chapter Four). A final possibility is to consider sufficient importance in terms of Gerrig's third and fourth requirements for suspense, namely that it is heightened by a challenging problem space and sustained over extended periods; both of these can be seen to have clear analogies in the musical domain. Some problems in a narrative have many solutions, whereas others have few; the same holds within a tonal musical idiom: for example, common intervals can imply many possible keys and melodic continuations whereas rare ones imply very few (Butler, 1989). Intervals that suggest a large number of possible keys and melodic continuations will tend to lead to the formulation of far more conflicting expectations and therefore, by definition, more uncertainty than those

[11] Of course, people regularly listen to music with which they are very familiar and they presumably do not find the repeated listening at all dull; this issue will be considered later on in the chapter.

that suggest but few; by extension, the problem space posed by these progressions will be somewhat wider. Conversely, intervals suggesting few obvious keys and continuations will lead to the generation of few expectations, but could maybe generate *suspense* nonetheless in a literal sense, as listeners wonder how the music can possibly continue (Narmour, 1992). In other words, it might be said to have a severely restricted *solution* space!

Music consists of multi-layered hierarchical structures within which expectations can build on multiple levels and multiple time courses. Notably, Lerdahl and Jackendoff (1983) have constructed a whole theory of western tonal music which takes this observation as its central tenet; more recently, this theory has been supported by extensive empirical work (e.g. Deliège, Mélen, Stammers, and Cross, 1996; Tillmann, Bigand, and Pineau, 1998). This potential for hierarchical, multi-layered uncertainties suggest a way in which musical "suspense" could be maintained over time: it is not difficult to conceive of music that perpetually sets up expectancies at all levels, ensuring that at any given time, at least one level remains in suspension. In the light of these observations, we can return to the issue of "sufficient importance". Just as a reader who becomes involved in a task of problem solving whilst engaging with a suspenseful narrative adopts the problem as his own, a listener engaged with the complex structures of a piece of music is generating expectancies of his own in response to that music. Just as the concept of empathy is not necessarily required, then, to explain a reader's personal involvement with a suspenseful text, neither is the concept required to explain why a listener might be moved by the generation, confirmation and violation of expectancies in music. A musical "outcome" is "sufficiently important" to generate suspense if the listener has been captively involved in generated expectancies which may be confirmed or confounded, or immersed in a problem space that is sufficiently complex.[12]

Theories of expectancy, such as that of Meyer (1956) have often been criticised for their highly Western-centric approach, leading to conclusions and explanations which cannot be

[12]The possibility of a relationship between complexity and emotion has been mooted several times elsewhere, most notably by Berlyne (1971) who modelled it as an inverted-U: for him, moderately complex music is maximally arousing; as the music becomes either more or less complex than this optimally moderate level, arousal decreases. The present conception of expectancy as narrative suspense provides an interesting framework for the interpretation of this notion: as music becomes more complex, it sucks the listener further and further into its problem space, leading to the establishment of "sufficient importance" and, hence, arousal; if, however, the music becomes so complex that it is no longer readily comprehensible, the listener will voluntarily withdraw further and further from the activity of problem solving, the requirement for sufficient importance will fail to be met, and level of arousal will decrease. A rather different, but not necessarily incompatible explanation for arousal in the context of problem spaces is suggested by an observation made by Bordwell (1985, p.39): "when we bet on a hypothesis, especially under the pressure of time, confirmation can carry an emotional kick".

generalised to other forms of music and culture (see e.g. Serafine, 1983). Unfortunately, no theory comparable to Meyer's but which focuses on another musical culture is currently available, so the question as to whether other musics contain structures that can fulfil the four requirements for suspense must remain beyond the scope of this thesis. Considering the generation of expectancies to be a peculiar type of narrative comprehension, however, renders this question somewhat uninteresting: rather than attempting to demonstrate that musical expectancies themselves are a pan-cultural phenomenon, we can merely hypothesize them to be one specific instance of such a phenomenon and, incidentally, one that can help account for the dynamic nature of listening experiences.[13] As Bruner (1986) and Shweder (1994) suggest, narrative represents a general and ubiquitous mode of human thought. From the perspective of the researcher interested in emotional response to music, adopting the stance that expectation is merely a form of suspense in the Gerrigian mould obviates altogether the necessity to treat it as a special—perhaps even natural—kind; in fact, expectation itself becomes utterly unimportant beyond its role as provider of an illustration of one way in which narrative experience can be evoked by music. A perpetual flux of tensions and resolutions may be at the heart of the dynamics of western tonal music, and may even be crucial to its affective power, but ironically, there may be nothing exclusively musical about it. Branigan (1992, p.38) wrote that "narratives are composed in order to reward, modify, frustrate, or defeat the perceivers' search for coherence"; one could imagine a similar claim being made for the dynamics of western tonal music. Expectancy is suspense; music is narrative.

Rehearing and Anomalous Suspense

As an explanation for emotional response to the dynamics of western tonal music, expectancy generation, violation and confirmation have always been confounded by one major snag: expectancies can only be violated for listeners who are experiencing a piece for the first time; someone who knows a piece well, and who therefore knows exactly what to expect, should only ever have expectations confirmed. This problem is not ameliorated by conceptualising expectancy as a specific instance of narrative suspense either because, in these terms, rehear-

[13]The phenomenon of expectancy as discussed here is not synonymous with the low-level, *gestalt*-type expectancies considered in Chapter Two. It was claimed in the earlier chapter that low-level expectancies are a function of our perceptual system and therefore pan-cultural by definition; the present discussion is not intended to supplant or contradict that claim.

ing a piece must be a suspense-free experience, as one of the primary requirements for the generation of suspense—the presence of uncertainty—is conspicuously absent.[14] Strangely, such a suggestion is totally contrary to common sense: people do listen to music repeatedly, and when they do so, they react to expectancy violations almost as though they are hearing them for the first time.

This problem is one that bothered Meyer (1961) considerably. Although he makes bolder claims than the present thesis for the importance of expectancy in music's power to evoke emotions, the essential dilemma is the same:

> "A theory of communication in which the unexpected, the ambiguous, and the less probable are of crucial importance for the understanding of, and response to, music is apparently in direct conflict with the belief that good music can be reheard and re-enjoyed countless times"(p.260).

Meyer (1961) provides several possible explanations for the apparent disparity, most of which are tightly focused on the perception of western art music, but do not completely resist generalisation. The first is that knowledge of the axioms of a piece of music are required before a listener can fully understand their implications. When hearing a piece for the first time, a listener can do little more than create a mental map of the main motifs, patterns and structures; only on repeated hearing will he or she be able to use these landmarks as a basis on which to build sets of expectations, implying that expectancies can only be generated by a listener who has at least some knowledge of the terrain. The second explanation is that listeners remember some parts of a piece of music rather better than others; even if only due to something as mundane as the limits of information processing ability, there is always something new to hear when rehearing a piece of music. It follows from this that there are always new expectancies to generate. A third suggestion is that musical predictions are more dependent upon a listener than on a work; the way in which the music is heard, and the expectancies that are generated, is dependent on a listener's state of mind and mode of engagement with it. Yet another suggestion is that each performance is a unique work of art, so in a literal sense every listening experience, even if it involves a familiar piece of music, is intrinsically a novel one. Such a suggestion, of course, can only be relevant to the hearing of

[14] The term "rehearing" is borrowed from Meyer (1961).

live performances, as every rehearing of a recording is a rehearing of the same performance as well!

The distinction between an initial exposure to a piece of music, where knowledge of a terrain is restricted to a listener's general cultural knowledge of the style (e.g. the tenets of western tonal music), and a second hearing in which the terrain expands to include aspects of the work itself, parallels the distinction that Bharucha and Todd (1989) draw between *schematic* versus *veridical* expectancies: schematic expectancies are those that are derived from a listener's knowledge of the style structures of a musical idiom, whereas veridical expectancies are generated in response to the note-by-note unfolding of a particular work. Conflict between schematic and veridical expectancies "underlies the tension between what one expects and what one hears"(Bharucha and Todd, 1989, p.44).[15] Bharucha and Todd do not share Meyer's view that the source of expectancy generation in reheard music lies in the fact that there is always something new to hear, even in a familiar piece; as far as they are concerned, there is *never* anything new to hear: listeners "know more than just what musical structures are likely in various contexts in their culture; they know exactly what event is to occur next to a particular point in particular pieces of music with which they are familiar"(p.50). Bharucha and Todd (1989) were interested primarily in the power of neural nets to model ways in which schematic and veridical expectancies might be acquired and developed—a subject of no direct relevance to present concerns—but one of their findings is of interest here: their neural net simulations found it harder to learn veridical expectancies that conflicted with schematic ones than ones which were congruent with them. This corroborates Meyer's suggestion that there is always something new to hear when listening even to familiar music; the novelty lies simply in the difficulty associated with embodying contra-schematic expectancies.

Although Meyer's suggestions as discussed so far have contained elements of plausibility, and although the work of Bharucha and Todd gives potential insight into possible causes of uncertainty when rehearing music, all these observations are very closely tied to the concept of western art music and listening processes peculiar to this music; as such, they all

[15]This distinction between schematic and veridical expectancies is not unique to Bharucha and Todd (1989); a similar distinction between cultural and specific knowledge is discussed by Dowling and Harwood (1986) who attribute the observation to Wittgenstein (1966). More generally, this parallels the distinction between so-called "episodic" and "semantic" memory, as discussed extensively in the psychological literature (see e.g. Tulving, 1993).

appear to be somewhat *ad hoc* and, as such, are not entirely convincing. Meyer makes one final suggestion regarding the possible source of expectancy violation and confirmation in reheard music, however, whose tenor is totally different:

> "Just as we are able to believe in ... the reality of a dramatic action ... even though we have seen it before and know what will take place, so too we are able to believe in the reality of a piece of music—to become involved in its syntactic structure—even though it has been heard before"(p.262)

This suggestion seems at first sight to be rather vague, and yet is remarkably persuasive. Put another way, it says that the paradox of finding the source of uncertainty in familiar music is exactly the same as the paradox of fiction, whereby suspense can be maintained throughout a novel or a film even though it may have been read or seen several times before. Once again, a phenomenon related to expectation in music is seen to bear a striking similarity to a phenomenon concerning suspense in narrative; in fact, the very words Meyer uses to draw his comparison alludes to a notion at the heart of this chapter, namely that the experience of listening to music is , at least in part, a narrative experience. Such a suggestion, of course, allows for the possibility that a consideration of *anomalous suspense*[16] in drama could provide us with another, rather different way of approaching the problem of expectancy in reheard music.

Psychologists studying the experience of narrative have often been fascinated by the phenomenon of anomalous suspense. For Walton (1990), it is symptomatic of a more general phenomenon whereby a narrative immerses the reader in a make-believe world into which previous real-world knowledge does not impinge. In this suggestion, Walton is not alone; Gerrig (1993, p.160) observes similarly that the main hallmark of anomalous suspense is that "what should be 'readily available' (by virtue of solid prior knowledge) is not, in fact, readily available in the experience of a narrative". Hence, prior knowledge of a story's resolution does not in any way detract from the suspense whilst re-reading it.

The suggestion that readers are immersed in a fictional world when they are engaged with a narrative and that this world, just like the real one, unfolds in a single temporal direction within which outcomes cannot possibly be known—that is, they are not available to consciousness at the time of re-reading—does not on its own seem farfetched. What is

[16] Terminology borrowed from Gerrig (1993).

rather peculiar, however, is the existence of abundant evidence suggesting that far from being sealed entities, the contents of narrative worlds has the power to affect people's judgements in other worlds, including the real one! Gerrig (1989) carried out a series of experiments to test the effect of anomalous suspense on reader's memory for real historical facts. He presented readers with short stories about commonly known historical facts, some of which were written in order to be suspenseful, and some which were not; half the suspenseful stories implied the factual outcome, and the other half implied a counter-factual outcome. Just before having had a chance to read the conclusions, readers were presented with a straightforward factual or counter-factual statement about the subject of the story: the task was to state whether the statement was true or false *in the real world*. Readers were consistently slower and less accurate in their answers having read suspenseful stories that implied a counter-factual outcome than in any other condition, even if they were told the correct answer immediately before reading the story. Gerrig concluded from this that participants considered the implied conclusions of the suspenseful stories even when being asked explicitly to make real-world judgements.[17]

Evidence that anomalous suspense is not restricted to engagement with fiction, and that that the salience of information in a narrative world is such that it even affects real-world judgements regardless of prior knowledge, suggests a narrative-based explanation for the generation of expectancies in reheard music. It seems plausible that each time a listener hears a piece of music, he or she might be immersed in a "narrative world" within which the "outcome" of musical sequences is unknown, regardless of the fact that it is known *in reality*. Therefore, within the narrative world of, say, a piece of music in simple tonal style, a subdominant followed by a dominant chord will imply a tonic resolution; a listener will always be surprised if these chords are actually followed by a submediant (an interrupted cadence) however many times the piece is heard. Further, if the piece is full of interrupted cadences, so that during the course of listening they become totally expected, this should have some effect on a listener's generic expectations, forcing interrupted cadences to become the expected norm. Confirmation of this hypothesis would require empirical research, but it fits comfortably with the relationship between schematic and veridical expectancies as elucidated by Bharucha and Todd in their neural net research.

[17] This series of experiments will be discussed further in Chapter Eight, which will report on a new experiment whose rationale was based on the Gerrig's methodology and findings.

Such a suggestion also fits well with the *parallel multiple-analysis model* of music perception posited by Jackendoff (1992). In this model, listening to music involves the generation of multiple parallel analyses of the abstract form that might be responsible for a given musical surface; "when the processor encounters a choice point among competing analyses, processing splits into simultaneous branches, each computing an analysis for one of the possibilities. When a particular branch drops below some threshold of plausibility, it is abandoned"(p.62). Within this model, each of the potential analyses will imply prospective expectancies, only some of which will be borne out in practice; in retrospect, the borne out expectancies will force a dynamic re-evaluation of the plausibility of previous analyses. Listening to music, then, is seen as a constant dynamic of competing alternative hypotheses as to the music's progression. Crucially, the processor responsible for inventing these analyses is *informationally encapsulated* in that it has "only the rules of musical grammar at its disposal to develop an analysis. Moreover, its operation is obligatory: in response to any plausibly musical signal, it tries its best to develop a musical structure"(p.67). In other words, this model posits that even in highly familiar music, the listening process would involve constructing these multiple analyses; our interrupted cadence, therefore, will *always* evoke the same reaction in a listener because "even if the listener consciously knows [about it] ... the processor [responsible for music listening] itself is innocent of this knowledge"(p.67).[18]

Despite the apparent ease with which analogies between expectation in reheard music and anomalous suspense in narrative can be made, and regardless of whether analysis of narrative comprehension might therefore offer a useful tool for the comprehension of reheard music, a further question remains: are the similarities merely metaphorical, or are the processes involved in the creation of these types of phenomena homologous? According to Gerrig (1993, p.170):

> "Anomalous suspense arises not because of some special strategic activity but rather as a natural consequence of the structure of cognitive processing .... Readers experience anomalous suspense because an *expectation of uniqueness* is incorporated within the cognitive processes that guide the experience of narratives".

He points out that in the real world, generic types of situations repeat themselves but spe-

[18]This model of music perception and the concept of a "processor" operating independently of a listener's conscious knowledge is, of course, reliant on a modular theory of mind (see Fodor, 1983).

cific instances do not; therefore it would be a total waste of cognitive resources to attempt to recall knowledge that is particular to a one-off situation (a particular novel, or a particular piece of music) in another, different situation; only common general heuristics should be extracted from one situation and re-applied in another.[19] When re-reading a novel, or re-hearing a piece of music, we are faced with a peculiarity in real-world terms: a situation that repeats itself exactly. If Gerrig's argument is correct, it follows that generating expectancies in reheard music should involve identical cognitive processes to those involved in generating suspense in re-read novels. Unfortunately, much empirical work would be required to ascertain the extent to which such processing is indeed generalisable, but "expectation of uniqueness" is a simple, persuasive concept. Gerrig's point is summed up by a neat aphorism that he quotes from Mason (1988):

> "Baseball is the same situations over and over, but no two turn out alike. Like crops and the weather. Life".

To Other Worlds

This chapter opened by introducing the seemingly ubiquitous phenomenon of music being perceived as narrative. It noted evidence of a stark contrast between the apparent effortlessness with which people listen to and discuss music as though it were narrative, and the immense difficulty facing theorists who attempt to specify what exactly a musical narrative might be. It was suggested, however, that from the perspective of a music psychologist, consideration of narratives in music not as objects, or properties of music itself, but as an emergent property of the listening process could provide a powerful tool for understanding the dynamic and coherent nature of the experience of listening to music. The second part of the chapter was devoted to a discussion of suspense (a well-documented narrative phenomenon) and its striking similarity to musical expectation, possibly the single most important contributor to the dynamics of emotional response to western music. But if music is heard as narrative, this cannot be simply because musical dynamics take a similar form to narrative dynamics; people do not hear music merely as narrative structure but also—as the extensive musicological literature on musical "meaning" and common sense conspire

[19] This line of argument is not unique. Similar claims regarding efficiency of cognitive resourcing have been made in a totally different context by e.g. Pinker (1997).

to reveal—as narrative *content*. To listen to music is to be immersed in an entire narrative world; to respond to music is to respond to events that are perceived to take place in that world. In this, the final part of the chapter, we will consider how listening to the dynamics of musical structure, context, sounds and utterances *as narrative* might impose coherence on a listening experience.

According to both Gerrig (1993) and Bruner (1986), central to the experience of narrative is the feeling of being "transported" to narrative worlds. Fundamentally, what makes these "worlds" distinct entities, and engagement with them identifiably distinct experiences is the fact that they have their own structural integrity and coherence, or to put it another way, "comprehending a narrative requires assigning some coherence" (Bordwell, 1985, p.33). How the comparatively few words in a text can create such rich and coherent structures has been a major question for researchers of narrative experience; the answer appears to lie in our ability to fill in missing information by inference and to apply both contextual and extra-contextual knowledge until coherence is achieved. From this perspective, Meyer's theory of expectation in western tonal music—that expectations are perpetually generated then violated or confirmed as the music unfolds—is essentially a theory of how the structure of western music helps a listener create a narrative experience: it explains how *knowledge* of a musical style can lead to the generation of *inferences* about the musical dynamic; these inferences provide the coherence required to maintain the existence of a narrative world. The reason why phenomena described by Meyer are so readily comparable with the phenomenon of suspense, as described by narrative psychologists, is that the experience of suspense, too, arises fundamentally due to a combination of knowledge and inference. As for how these structural inferences fuse with narrative content, a clue comes once again from Bordwell (1985, p.49), when he writes:

> "as a dynamic process, narration deploys the materials and procedures of each medium for its ends .... Narrative patterning is a major part of the process by which we group films as more or less coherent wholes".

In other words, anything and everything that is perceived as part of the same whole can be fused into a single narrative.

An alternative, but not at all incompatible explanation for coherence is proposed by Lakoff and Johnson (1980), who claim that crucial to the coherent structuring of experience is the idea of metaphor. According to them, far from being a specialist literary device as might initially be supposed, metaphor is fundamental to the way we conceptualise the world, and "as much a part of our functioning as our sense of touch"(p.239). Their book is replete with examples of metaphors embedded deep within in our conceptualisations of everyday life. For example, they devote a considerable time to the metaphor ARGUMENT IS WAR, noting that not only is the structure of an argument based upon the structure of war—it involves attack, defence, retreat, manouveuring, followed by stalemate, truce, surrender or victory—but also arguments are *understood* in terms of war, affecting profoundly the way in which we argue and the definition of argument itself. Or as Lakoff and Johnson themselves put it:

> "one activity, talking, is understood in terms of another, physical fighting. Structuring our experience in terms of such multidimensional gestalts is what makes our experience *coherent*"(p.81).

By way of explanation for this phenomenon, Lakoff and Johnson note that "we need to classify our experiences in order to comprehend"(p.85); metaphor allows us make sense of the world in terms of structures that we know and understand.[20] In the light of these observations, we can perhaps suggest that music is heard as narrative because when we listen to music we conceptualise it *in terms of* narrative, with narrative itself acting almost as a meta-metaphor within which all things can be made comprehensible. Structured as narrative, listening context (associations, environment, mood), sound, utterance and the moments of physiological arousal that they can evoke all cohere into a unified but dynamic experience.

If the similarity between phenomena found in the experience of reading and the experience of listening to music can be attributed to common cognitive processes, and if the rationale for the commonality can be found in the common creation of coherent narrative worlds, then two theoretical issues can be resolved. First, we can provide a further explanation for the apparent lack of generalisability of musicological theories such as that of Meyer (1956):[21] the problem is not so much that the scope of the *theory* is limited to the explanation of western tonal music, but that the structures upon which it focuses *only exist* in western tonal music. It has often been observed that the haloed status we give to music, particularly art music,

[20]Unfortunately there is no space here for a further elucidation of Lakoff and Johnsons' extensive theory or any more of the literally hundreds of examples of metaphors that they present; the interested reader is strongly encouraged to read the original book.

[21]See Chapter One of this thesis.

is a western peculiarity;[22] this peculiarity is essentially that we seek meaning—narrative content—in the music itself, rather than, say, in an occasion or in a text. By contrast, it has always been clear that in cultures where music plays a specific social or ritual role, narrative content inheres in the event rather than the notes, and that in these cultures, the knowledge and inferences required to create a narrative world presumably do not relate directly to the music. As we have now seen, however, even in western music Meyeresque expectations cannot by any means take sole responsibility for the creation of narrative worlds; they merely provide the dynamics upon which narrative processes can feed.

The second conclusion to be drawn from all this is that to continue the debate as to the ontological nature of musical narratives—as elucidated in the first half of this chapter—would be a vain pursuit indeed. Humans possess a pre-disposition to try to make sense of the environment around them; when faced with music, they construct narrative worlds to impose a coherent structure on their experiences. Construction might be based on perception of musical parameters wholly, partly, or not at all; the balance is likely to vary considerably from culture from culture, from piece to piece, from person to person and from context to context. As for emotional responses, they might be evoked as a result of a listener being unable to construct a coherent narrative or, more usually, as a result of the implications of a narrative experience. Either way, it seems impossible to deny that emotional response to music is, at least in part, emotional response to narrative, simply because when people listen to music, they are listening to it *as* narrative; if nothing else, hearing as narrative can certainly account for the sense of coherence experienced by listeners who respond emotionally to music. After all, as Branigan (1992, p.29) wrote in a sentence that mirrors Bruner's (1986) description of the *narrative mode* cited near the start of this chapter, "a narrative schema, together with a host of related schemas, encapsulates the interest we take in the world as humans". To claim that we hear music as narrative, then, is tantamount to claiming that when we listen to music we try to integrate sound, utterance and context into a coherent structure with a dynamic "cause and effect teleology"(Branigan, 1992, p.19). It is for this reason that a suggestion that music is heard as narrative constitutes the fourth and final premise of the model of emotional response to music promoted in this thesis.

[22] Although Bohlman (2000) notes that it is only thanks to the recent influences of ethnomusicology and music sociology that musicologists have appreciated that this is so.

Chapter 6

The Model and its Implications

The previous four chapters described the four components of a new listener-centric model of emotional response to music, and discussed the rationale for their formulation. This, the final chapter of the first part of the thesis, presents a summary of these four components, concentrating particularly on describing interactions between them. Then, it considers the model as a whole in terms of the need outlined in Chapter One for a theoretical framework within which empirical research on emotional response to music can be understood.

However complex a musical structure and however multi-faceted a listening experience, we cannot avoid hearing and responding to music as sound: acoustic signals of certain intensity or timbre cause physiological arousal necessarily and unavoidably, without conscious intervention, owing merely to the low-level workings of our perceptual system and the evolution of survival-oriented parts of the brain. Likewise, groupings of acoustic stimuli form patterns that generate perceptual expectancies regarding their continuation; when such patterns violate the expectations, physiological arousal also ensues. Again this process occurs without any conscious intervention; it is simply a function of perceptual and emotion systems that are optimised to detect and respond quickly to unexpected changes or uncertainties in our environment. So low-level and ingrained are these responses that they cannot be turned off at will when we stop listening to the sounds of our environment and turn instead to the sounds of music; rather, they become an integrated part of the musical listening experience.

As Darwin (1872) and many commentators since have noted, we are remarkably good at detecting the emotional state of fellow humans from their facial and vocal expressions.

Furthermore, when we detect such expressions, we cannot help but respond to them: a blood-curdling scream, a quiet whimper or a shout of elation inevitably lead to physiological arousal in a listener. It appears that immediate empathetic physiological response to human emotional signals is a highly ingrained, pre-conscious phenomenon; sensitivity to emotional utterances and gestures is apparent even in newborn babies. Many musical sounds are literally vocal in that they centre around the singing voice; listening to such music involves detecting and responding to emotional cues encoded within the vocal line just as much as does listening to a human voice in a non-musical context. In addition, there is strong evidence to suggest that an acoustic signal that sounds plausibly like a vocal utterance, even if actually produced by a musical instrument, is detected and responded to as though it were indeed such an utterance. Of course, conscious introspection and high-level interaction with the music would make ultimate diagnosis and interpretation of such a sound somewhat different from it's vocal counterpart; for that matter, there would also be a vast difference on this level between a vocal sound produced as part of a musical performance and one produced as a result of a real-world situation. Despite the possibility for differentiation via these high-level cogitations, however, we cannot avoid responding immediately and automatically to something that sounds like a human emotive utterance. Thus, hearing music as utterance—perhaps also as gesture—is an important contributor to the emotionality of a listening experience.

Responding emotionally to music is not simply a matter of reacting helplessly to sounds that trigger survival-instinct type emotions and to those that sound like emotive human utterances. It also involves responding to the flood of thoughts and associations that can be sparked by a listening experience owing to the listening situation, autobiographical memory, musical or poietic knowledge; it involves responding to context. Music provides a rich context for cognition, sparking associations—musical and extra-musical—that could be related to any aspect of the listening situation; once sparked, such associations can evoke emotions and mediate mood through their influence on cognitive appraisal mechanisms. Crucially, once sparked, they and their emotional effects become an integrated and salient part of the listening experience.

When we listen to music, we attempt to hear the notes, the musical structures and all salient aspects of the listening experience as a dynamic, coherent whole; such an activity can be described as listening to music as narrative. A narrative experience can be emotional if it is suspenseful or violates expectations; alternatively, emotion can derive from interpretation of the narrative content, that is, the moment by moment implications of a constantly unfolding context. Unlike sound, utterance and context, hardly any empirical work on emotional response to music has considered the experience of listening to music as a form of narrative experience, but the wealth of evidence from other domains, such as literature and film suggests that far from being a domain-specific phenomenon, narrative is fundamental to the way in which we understand the world. In every domain, "the human species is compelled ... to make sense of both life and the universe"(Storr, 1992, p.105); there is no reason to believe that listening to music is any different.[1]

Sound, utterance, context and narrative are not posited here as independent, isolated phenomena or as convenient buckets within which certain classes of empirically observed phenomena can be grouped. By contrast, it is proposed these four aspects of a listening experience and, perhaps more crucially, interactions between them are the basis of a coherent—if generic—model of emotional response to music that is capable of embracing all reported phenomena. A graphical representation of this model is presented in Figure 6.1, illustrating interactions between sound, utterance, context and narrative, and their relationship to physiological arousal as evidenced by the previous four chapters.

At the base of the model sit sound and utterance. Both can be seen to have a direct effect on physiological arousal (indicated by arrows I and L), reflecting that hearing music as sound and music as utterance can cause arousal directly, without cognitive mediation. The lack of any incoming arrows pointing at either sound or utterance also reflects their comparative independence or robustness; the suggestion is that emotions evoked by these modes of listening are so low-level and automatic that they cannot be ameliorated by any other aspect of listening. The gunshot or the anguished cry are no less potent for being part of a musical listening experience rather than part of real life—not, at least, at this level. Both sound and utterance also have an influence on context (arrows G and J) and on narrative (arrows H and K): although immediate response to certain sounds and utterances may be

[1]Storr (1992) is one of the few theorists who comes very close to invoking narrative as an explanation for emotional response to music when he writes: "I am sure that one of the reasons why music affects us deeply is its power to structure our auditory experience and thus to make sense out of it"(p.105). He does not mention narrative by name, however; sadly, perhaps owing to the nature and size of his book, he does not develop this idea further.

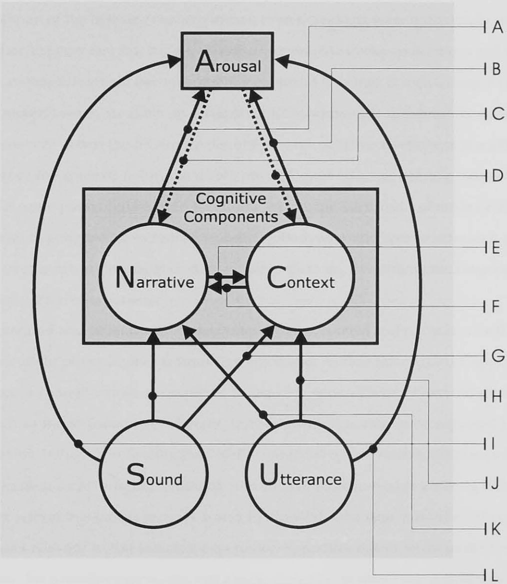

Figure 6.1: Sound, Utterance, Context and Narrative – the complete model
Sound, utterance, context and narrative each affect level of physiological arousal (I,L,B,A); sound and utterance influence context and narrative (G,J,H,K); narrative and context are interdependent (E,F); arousal feeds back and influences context and narrative (D,C). Context and narrative are the cognitive components which, together with arousal, constitute emotional response to music. Note: sound, utterance, context and narrative here are attributes of listening, not attributes of music.

direct, all such sounds and utterances are open to conscious interpretation, and thus impinge on the context in which music is heard, and the way in which narrative coheres. A gunshot in a crowded street may just be a loud bang on one level, but it is also *recognisable* bang, whose presence is replete with associations and valenced concepts that, once brought to mind, form an integral part of the listening context; in fact, even if the bang were not actually recognised as a gunshot, the very fact that it is a loud noise is enough to evoke associations and cognitive correlates of fear that would become part of a context. Of course, if such a bang were clearly part of a musical work, its effect on context could be somewhat different; the crucial thing to note, however, is that the difference lies in the way cognitive assessment of a loud sound forms part of the context, not in the direct physiological effect of hearing the sound itself. Exactly the same points can be cited for utterance. As for the direct influence of sound and utterance on hearing music as narrative, hearing both as sound and as utterance generates a flow of events of varying salience that we can expect might become marker points in the structure of narrative experience.

Above sound and utterance in the representation sit context and narrative, each of which can contribute to physiological arousal (arrows B and A) through cognitive appraisal: the main effect of context would normally be to mediate general level of arousal through processes such as mood induction, whereas narrative would normally be expected to exert a more dynamic influence, reflecting the volatile nature of narrative-structured experience. In addition to their direct effects on arousal, narrative and context have a strong influence on each other within this model (arrows E and F). Context—the sum total of all thoughts and propositions relevant to the listening experience—provides not only the general cognitive background for narrative processing, but also moment-by-moment nuggets of stimulus that form the *content* around which narrative is structured. A concept of closure from a cadential progression, or recognition of the *Death* motif in Mahler's *Symphony No. 6* might both emerge from listening to music as context; they would also be expected to impinge on listening to music as narrative. Conversely, the state of an unfolding narrative—its tensions, its unfolding structures, its inevitabilities and the extent to which the listener can make sense of all of these—become in themselves part of the context in which music is heard; narrative influences context.

As we have already seen, both context and narrative are influenced by sound and utterance (incoming arrows G,J,H and K); each is also affected by feedback from arousal itself (indicated by the dotted arrows D and C). A listener's state of physiological arousal is as much part of the listening context as anything else, and possibly a more salient part than most! As for the feedback effect of arousal on narrative, we should expect that general level of arousal will affect the dynamics of narrative formation largely by mediating the apparent "importance of a target outcome", to use Gerrig's (1993) terminology; in addition, such feedback could affect the apparent "speed" or level of activity of narrative formation.

In the graphical representation of the model, context and narrative are enclosed within a gray box and labelled as "cognitive components", reflecting the definition of these two levels of listening as consciously mediated, susceptible to external influences, and available to introspection. The enclosure, however, has an additional significance: there are only two rectangular boxes in the diagram, this one surrounding context and narrative, and one representing physiological arousal; between them, these two boxes encapsulate what this model identifies as the *emotion* in emotional response to music. After Schachter and Singer (1962), emotion is conceptualised as physiological arousal coupled with cognitive explanation of that arousal.[2]

There are several things to note about this model. First, as has already been stated, it claims to be *complete* but *generic*: it neither claims nor even aims to provide a full account of how emotion is evoked in a particular listener experiencing a specific piece of music in a specific situation; instead, it suggests how the almost infinite array of parameters that might possibly impinge on such a process can be understood as a coherent whole rather than a collage of disparate phenomena. Listening to music involves a vast number of highly complex processes, many of which are only partially understood; the task of a model of emotional response to music, surely, is not to attempt full explanation of all of these, but rather to show how an eventual understanding of them would relate to emotional response. The model itself also refrains from positing its own emotion theory in order to account for emotional response to music: the combination of arousal and cognitive components outlined above constitute a *description* of what is meant by "emotion" in the phrase "emotional response to music", but in no way suggest a specific *mechanism* by which this emotion may occur. Far from being a limitation, this is one of the model's most powerful features: the evocation of

[2]Of course, according to the model here, the cognitive components also partly *cause* the physiological arousal; but that is an entirely separate issue!

emotions is a subject that has sparked significant theoretical and empirical research in its own right; rather than vainly attempting to displace such work with a model of emotion moulded conveniently to explain phenomena associated with emotional response to music, this model instead strives to account for emotional response to music within the ambit of existing models of emotion.

In the first chapter of this thesis, one of the criticisms made of the Dowling and Harwood (1986) model concerned its categorisations of phenomena which, it was claimed, are somewhat arbitrary and do not well reflect empirical reality. It should be noted that the model proposed in this thesis also categorises phenomena, but the distinctions that it draws are qualitatively very different in nature. First, the categories of sound, utterance, context and narrative fit well with existing empirical data, an observation that is perhaps truistic given that the model emerged from a consideration of such data. Second, as has been stressed throughout this thesis, the categories here distinguish between modes of listening rather than aspects of a musical environment: the model does not claim that music *consists* of sound, utterance, context and narrative; rather it claims that music *is heard as* sound, utterance, context and narrative. The gulf between these two propositions is not merely a question of semantic pedantry; rather, the listener-centric distinctions being made in the present model are entirely incompatible with the alternative artifactual categorisations. For example, a piercing shriek could most definitely be defined as an "utterance object"; in the present model, however, perception of that utterance would not be classified solely as "hearing music as utterance", but would impinge on the other three domains as well. It would certainly be heard as a sound, would invoke plenty of contextual associations and could constitute a significant element of dynamic narrative process. A more clearly musical artifact, such as an interrupted cadence, would likewise impinge on emotional response in all four domains. In short, the present model is not so much optimised for classifying musical phenomena but rather for understanding the way in which perceptual and cognitive phenomena might lead to musically-evoked emotion.

A result of the present model's focus on the activities of a listener and its avoidance of object-based categorisations is that it draws no distinction whatsoever between intra- and extra-musical phenomena; the previous four chapters of this thesis have, in fact, taken pains to highlight the arbitrariness of such distinctions by pointing to the great equivalences between intra- and extra-musical emotion catalysts that can be observed in much of the empirical data, an enterprise that stands in stark contrast to the traditional pursuit of trying to find differences. Of course, the avoidance of these distinctions is crucial to the model's ability to take the stance it does concerning emotion, namely, that emotional response to music can best be understood in terms of what we know about emotion rather than what we know about music. As was suggested in the first chapter of this thesis, the fact that music psychologists have tended to draw distinctions between intra- and extra-musical phenomena can perhaps be understood in terms of the musicological heritage of music psychology as a discipline. Ironically, however, it is not just music psychologists who have assumed that music is somehow different from other things in the world and, by extension, that listening to music is different from indulging in every other activity; the notion has even spread to some students of emotion psychology, who have seen the existence of music as problematic for extant theories of emotion. For example, appraisal theorist Ellsworth (1994, p.196) writes:

> "Music ... [is] problematical for appraisal theories of emotion, but no other theories have done much better. The usual tactic is simply to omit any reference to these embarrassing mysteries, to rule out domains where our theories falter as irrelevant. .... Nonetheless, the mysteries *are* the theoretical challenges, and sooner or later must be acknowledged".

Even if it serves no other function, the present model and the empirical evidence that supports it at least challenge assertions such as this!

The first chapter of this thesis suggested that despite the impressive quantity of robust empirical research on emotional response to music, music psychology's empirical attempts to grapple with the subject were in serious need of a theoretical framework within which they could be understood and research paradigms developed. It is here claimed that through its lister-centricity, its avoidance of drawing arbitrary distinctions between intra- and extra-musical phenomena, and its compatibility with current emotion theory, the model presented here constitutes one such framework. By far the largest volume of research in the field has concentrated on understanding accultured listeners' responses to the structures and idioms of western art music; consideration of such responses as response to sound, utterance, context and narrative offers the possibility of generalising those findings and making cross-cultural or even cross-domain comparisons. In addition, it suggests a rationale for why certain structures appear to elicit similar responses in similarly acculturated individuals in similar situations; in other words, it can account for intersubjectivity. Perhaps more importantly, the model makes clear suggestions for a future research agenda. Implicit in the conceptualisation of emotional response to music as response to sound, utterance, context and narrative is the possibility that an understanding of emotional response to music will derive from an understanding of response to sound, utterance, context and narrative; or to reverse the sense somewhat, an understanding of emotional response to sound, utterance, context and narrative is all that is needed to understand emotional response to music. This notion will be considered further in the final chapter of this thesis.

As the second chapter of this thesis has illustrated, we already know a significant amount about emotional response to sound, albeit largely in a non-musical context; it is gratifying to see that despite the comparative dearth of research in this subject from a music psychological perspective, some pockets of interest have recently begun to emerge. For example, Balkwill and Thompson (1999) and Bhatti and Gregory (2000) have carried out experiments that have demonstrated a level at which response to music is not reliant on familiarity with a musical culture but rather depends on our response to sound itself. This research project was concerned with non-western music, and similar attitudes are not so apparent in the mainstream literature; its existence is nonetheless encouraging. Much work on emotional response to sounds has concentrated on loud or particularly violent sounds; although such sounds do appear regularly in music, one important avenue of future research for music psychology, surely, would be an investigation of the emotion-evoking potential of the many other types of sounds that form a staple part of music. This quest could be helped enormously by music psychologists willing to adopt the significant battery of experiment paradigms and theory that have been developed by auditory researchers studying this field.

Musicologists and philosophers have been likening music to human utterance and gesture for centuries, as the third chapter of this thesis has illustrated; there is also a wealth of empirical music psychological work that points to the relationship between music, utterance and gesture. Much of the empirical work, however, concentrates on production rather than perception, and because so little of it has been concerned explicitly with understanding emotional response to music, being pre-occupied instead with musical representation of emotion, there are many areas of research—such as investigation of expressive timing deviations in performance, and timbral characterisations of emotion amongst instrumental players—that could surely be seen as aspects of the same phenomenon, but are instead studied separately. According to the model presented in this thesis, there would be much to be gained from looking explicitly at perception of and response to human utterance and gesture in music, drawing both theoretically and methodologically on the extensive literature on infant's perception of human utterance, and on perception of and response to vocal affect.

A wealth of research, some reviewed in the fourth chapter of this thesis, has pointed to the hugely important effect of context on emotional response to music. Although by far the largest proportion of this research has come from a social psychological rather than a music psychological perspective, hardly anyone would deny the importance of so-called contextual phenomena on musically-evoked emotion. As has been discussed at length in the earlier chapter, however, many music psychologists, whilst accepting the importance of such phenomena, see them as something quite distinct from emotional response to music itself. Ironically, it has taken a musicologist—Cook (1998)—to elucidate the intuitively obvious fact that music and context are not separate entities, but a highly complex, intertwined whole. The present model of emotional response to music suggests a rather more integrated research agenda, that would attempt to investigate both aspects of musical structure and the distinctly "extra-musical", employing the same methods and approach.

The implications of considering emotional response to music as emotional response to narrative are rather less obvious than those emerging from the other components of the model. Clear analogies between aspects of emotional response to music and emotional response to narrative, as discussed in chapter five, suggest that emotional response to music could be at least partly understood as response to a form of narrative experience. Implicit in this suggestion is the proposition that an extremely valuable research agenda for music psychologists studying emotion would be investigation of the extent to which response to music can be understood in terms of what we know about narrative experience in general. An oft-noted characteristic of musically-evoked emotional experiences is that they tend to be dynamic and yet coherent; particularly exciting is the possibility that empirical investigation of music as narrative could account for this characteristic. Unfortunately, however, there is a major snag: whereas empirical paths for researching response to music as sound, utterance and context are trivial to find, owing to the profusion of existing research paradigms and methodologies that can be used and extended, no such obvious paths exist for investigating response to music as narrative. To be sure, a wealth of research on musical expectancy may be a starting point, but it is hard to see how such research can be generalised. A more profitable solution might be to borrow methodologies from work that has been carried out on response to narrative in other domains such as film or written text; after all, if the construction of narratives is a domain-independent human activity—a view which this thesis adopts—it should be expected that similar empirical methods in different domains could be used to reveal similar responses. Sadly, this idea must remain pure conjecture because such an approach seems never to have been tried! No evidence can be cited here to show that experiment paradigms used in other domains can be adopted or adapted for use in investigating emotional response to music as narrative. The experiment reported in Chapter Eight of this thesis attempts to go a small way towards rectifying this situation.

Narrative is not only a core component of the model of emotional response to music proposed in this thesis, but also probably the component with the most potentially far-reaching implications for research, owing to its suggestion that the coherence, dynamics and even complexity of musically-evoked emotions need not be observed as a unique human mystery but can instead be understood as a specific instance of a domain-independent human activity. Although the years of work required fully to substantiate such a claim lie far beyond the scope of this thesis, presenting some corroboratory evidence does not. The second part of the thesis will present three experiments that constitute a preliminary investigation of emotional response to music as narrative. They will use real musical excerpts presented in multimedia contexts with the intention of evidencing the involvement of narrative process in music listening and the emotions that it can evoke. In so doing, they will attempt to show that such investigations are not only possible, but also useful, able to shed new light on poorly understood phenomena. Finally, they will try to provide further evidence of the central mantra of this thesis: emotional response can best be understood not in terms of music, but in terms of a listener; a satisfactory explanation of the undeniable power of music to move can only truly be achieved through treating the experience of listening to music as though it were, as in fact it is, a perfectly ordinary human engagement.

Part II

An Empirical Investigation of Narrative

Chapter 7

Music in the Presence of Explicit Narratives

It was suggested at the end of Chapter Six that one of the most important and least well supported claims of this thesis is that emotional response to music is, at least in part, emotional response to narrative; that the process of listening to music involves detecting and interpreting cues from sound, from human utterance and from context—both circumstantial and musical—and binding these together into dynamic, coherent narrative structures. The claim is important because, if ultimately validated, it will allow our understanding of narrative structures in other, non-musical domains to provide an explanation for much of the phenomenology associated with emotional response to music; it is the least well supported because virtually no music psychological research appears ever to have been carried out explicitly to investigate the notion that emotional response to music involves narrative processes. Chapter Eight will use a methodology adapted from a non-musical study of narrative processes in an attempt to demonstrate equivalences between these processes and music listening. This chapter, however, presents an experiment that, using conventional music psychological techniques, tries to provide some preliminary evidence for the influence of narrative process in the generation of emotional responses to music. It does not by any means attempt to offer conclusive a study of the phenomenon—something that would be way beyond the scope of this thesis—but it hopes to illustrate that the investigation of narrative is a useful, valuable pursuit for music psychologists studying emotion.[1]

[1]Results of this experiment have previously been presented at the *6th International Conference on Music Perception and Cognition*, Keele, August 2000.

An obvious difficulty facing an empirical attempt to investigate potential narrative processes in music listening is that narrative processes—in common with many phenomena that psychologists investigate—cannot be observed directly, and further—unlike many phenomena that psychologists investigate—are likely to be highly resistant to introspection: one could hardly ask experiment participants to give an account of their experience of narrative formation, or to rate its impact on their emotional response using a seven-point scale! A plausible workaround might involve observing the outcome of a hypothesized narrative process and comparing observations with hypothesis-inspired predictions. Whilst such an approach is not at all alien to music psychology, it still leaves us with a problem in this case: the effect of narrative on emotional response is hypothesized to be dependent upon the elements that form a part of the constructed narrative; these elements derive not only from the music being presented, but also from the context, much of which will be idiosyncratic to a particular listener. It would not be possible, therefore, to make any firm predictions against which to measure results of a study. In order to resolve this difficulty, the experiment presented here attempted to restrict the potential for idiosyncrasy of context by forcing listeners to respond to music in the presence of rich and structured explicit narratives, *i.e.* stories. Although response to stories, just like response to music, may well vary idiosyncratically from person to person, it was felt that their propositional nature and the likelihood of shared cultural interpretation should serve greatly to restrict the degree of variation between listening contexts.

In the experiment, participants were asked to listen to a series of musical extracts whilst reading short stories; after each story-music complex, they were asked to assess on a set of seven-point rating scales their emotional responses to both music and story; ratings were compared with independently obtained measures of the emotion-evoking properties of the music and stories when presented separately, using both ANOVA and Regression-based statistical techniques. The experiment was based on the Gerrigian premise that narrative processes bind salient aspects of an experience into a coherent, integrated structure. To the extent that listening to music is a narrative process, it should involve the construction of a coherent structure that accounts, amongst other things, for salient features of the listening context; if central to the listening context is a story—a complete coherent narrative in itself—it should be expected that this narrative would, at least to an extent, become integrated into the "musical narrative". As a result, listeners' emotional responses as attributed to the music should be markedly affected by the emotional valence of propositional content of the narrative.[2] If, on the other hand, music is not heard as narrative, emotional response as attributed to the music should not change as a direct result of the valence of the story. Of course, it is quite likely that highly moving story could moderate mood which, in turn, would affect emotional response to the music, but participants should be able to factor this out when giving their ratings.[3] More specifically then, it is hypothesized that the stories will affect the valence of emotional response to music but not the level of arousal. Thus the experiment is simple—perhaps even crude—but hopefully serves as a useful probe of narrative involvement in music listening; further, it complies with a central theme of this thesis, namely that music psychologists studying emotion should consider music from the perspective of a listener, in a rich multimedia context.

The experiment actually consisted of two separate parts. In the first, a series of participants was invited to listen to each of the musical extracts and rate their responses to them on a number of dimensions. The purpose of this part of the experiment was two-fold: to test the efficacy of the experiment paradigm, checking that participants were able to use the rating scales successfully and that the stimuli chosen elicited a suitably wide range of responses; second, to provide a profile of the stimuli for use in Part Two of the experiment. In Part Two, a different set of participants was asked to rate responses to music extracts and stories when presented simultaneously; this was the main part of the experiment as discussed above. It was considered desirable to ensure that each participant read each story only once during an experiment session; unfortunately, this desire precluded the possibility of a Cartesian Product design where each extract was paired with each story, allowing easy comparison of the effects of different stories on the musical extracts;[4] Therefore, ratings from Part One of the experiment were used to ensure that whilst participants only read each story once, each was presented with music-story pairings capable of evoking a wide range of responses. The approach taken is described in detail below.

[2]The converse does not necessarily hold because the story is likely to be coherent enough in itself that it has no need to incorporate aspects of the music into its structure.

[3]Even if they are unable to do so, cross-contamination from context effects should affect level of arousal, but not the valence of reported response to the music.

[4]Time constraints also precluded this possibility: each experiment session would have to last several hours if a full Cartesian Product of trials were presented!

Part One Method

Participants were invited to listen to twelve musical extracts, and to answer the following five questions immediately after hearing each:

1. *How moving did you find the musical extract? (on a scale of 1 to 7, where 1 = as unmoving as any that I have heard; 7 = as moving as any that I have heard).*

2. *If you found the music moving, was it positively moving, negatively moving, or neither (positive = e.g. made me happy, amused; negative = e.g. made me sad, morose, contemplative; neither = I answered "1" to the previous question).* A fourth possible answer was *"don't know".*

3. *How much did you enjoy listening to this music (on a scale of 1 to 7 where 1 = as little as I have ever enjoyed listening; 7 = as much as I have ever enjoyed listening).*

4. *Did you think this was a good piece of music? (1 = as bad as any music that I have heard; 7 = as good as any music that I have heard).*

5. *Have you heard this music before?*

In addition, they were asked to read twelve short stories, and to answer the following equivalent questions after reading each:

1. *How moving did you find the story? (on a scale of 1 to 7, where 1 = as unmoving as any that I have read; 7 = as moving as any that I have read).*

2. *If you found the story moving, was it positively moving, negatively moving, or neither (positive = e.g. made me happy, amused; negative = e.g. made me sad, morose, contemplative; neither = I answered "1" to the previous question).* A fourth possible answer was *"don't know".*

3. *How much did you enjoy reading this story (on a scale of 1 to 7 where 1 = as little as I have ever enjoyed reading; 7 = as much as I have ever enjoyed reading).*

4. *Did you think this was a good story? (1 = as bad as any story that I have read; 7 = as good as any story that I have read).*

5. *Have you read this story before?*

Listening to all twelve musical extracts took just over 1 hour; reading all of the stories took between 45 minutes and 1.5 hours, depending on the participant's reading speed. Half of the participants performed the reading task first; half performed the listening task first.

Questions

The most important questions from the perspective of the present hypothesis were the first two in each group, concerning "movingness" and "valence"; taken together, these questions constituted a measurement of emotional response to the music or story.[5] The first question was deliberately phrased so as to encourage participants to assess how the music or story affected *them*, rather than eliciting a judgement of the emotional effect intended by a composer or writer, or emotional effect encoded within music or story as defined by cultural convention. In addition, the phrasing "how moving did you find the music / story" as opposed to say, "how much were you moved by the music / story" was intended to allow participants to compensate for their own contextual environment, for example, by allowing for an unusually good or bad mood; in other words, it encouraged participants to be intelligently introspective about *themselves* rather than analytical about the music or story in question. Informal comments from participants indicated that it was indeed understood as intended.

The "don't know" option provided in the second question was intended to cater for the possibility that whilst a listener may have been moved by the music, or a reader by the text, emotional valence may not have always been unambiguous.

The primary purpose of questions 3 and 4 in each group, which explicitly asked for an objective assessment of the merits of each musical extract or story, was to offset and thereby further to clarify the meaning of the "movingness" and "valence" questions.

The final question in each group was included for completeness; its meaning and purpose are self-explanatory.

Stimuli: music

Several criteria determined the choice of musical extracts: they should be from purely instrumental music, containing no obvious programmatic content or conceptual representation; they should be accessible to participants, and thus in a reasonably familiar idiom; between

[5]These two dimensions have been found in several studies to constitute an intuitive, reasonably reliable measure of emotional response (notably Russell, 1979).

them, they should be have the potential to evoke emotional responses in the listener at both ends of the valence spectrum (positive and negative).

In order to avoid making a completely arbitrary or idiosyncratic choice of extracts, all were selected either from musical works whose typical emotional effects have been documented in previous experimental work (specifically from Sloboda, 1991; Giomo, 1993; Peretz, Gagnon, and Bouchard, 1998).[6] After eliminating all vocal music from the list of possibilities, it was decided to choose pieces that were timbrally as similar as possible to each other, in an attempt to minimize possibly overpowering effects of Sound or Utterance being primarily responsible for differentiation of participants' ratings. This requirement, together with a wish to provide a rich and realistic musical experience for participants, led to all extracts being taken from orchestral music. The list of chosen works and their origins are listed below:

1. Bach, J.S. - *Orchestral Suite No.3, 3rd movement.* Stylistically similar to music rated as soft, pleasant and trivial / non-solemn (Giomo, 1993).

2. Beethoven - *Piano Concerto No.4 in G major, 2nd movement.* Sections have been reported to elicit tears and lumps in the throat (Sloboda, 1991).

3. Beethoven - *Symphony No.3 in E flat, 3rd movement.* Cited as example of happy music (Peretz et al., 1998).

4. Brahms - *Piano Concerto No.1 in D minor, 2nd movement.* Sections have been reported to elicit tears and lumps in the throat (Sloboda, 1991).

5. Handel - *Water Music Suite; 2nd movement.* Rated as soft, pleasant, and trivial / non-solemn (Giomo, 1993).

6. Mahler - Symphony No.3, 1st movement. Sections have been reported to elicit shivers down the spine of goose pimples (Sloboda, 1991).

7. Mozart - *Piano Concerto No. 23 in A major, 3rd movement.* Cited as example of happy music (Peretz et al., 1998).

[6]A single exception was Shostakovich, *Festive Overture*, which was chosen by the experimenter to fill what he considered to be a gap for joyful, loud, large-scale, triumphant music.

8. Rachmaninov - Piano Concerto No. 2 in C minor, 2nd movement. Sections have been reported to elicit shivers down the spine and goose pimples (Sloboda, 1991); cited as example of sad music (Peretz et al., 1998).

9. Rachmaninov - *Symphony No.2 in E minor, 3rd movement*. Sections have been reported to elicit tears and lumps in the throat (Sloboda, 1991).

10. Schoenberg - *Verklarte Nacht*. Sections have been reported to elicit shivers down the spine and goose pimples (Sloboda, 1991).

11. Shostakovich - Festive Overture Op. 96. Vibrant, joyful music chosen by experimenter.

12. Sibelius - *Symphony No. 7*. Sections have been reported to elicit tears and lumps in the throat (Sloboda, 1991).

It was required that each musical extract should sound for approximately the length of time taken for an average reader to read one of the short stories; therefore, it was decided that each should last approximately 5 minutes. Extracts were created from commercially available CD recordings of each of the chosen musical works, using the following procedure: where the movement in question was approximately five minutes long (the case with the Bach and Handel) these were copied directly on to a master CD; where it was substantially longer than this (the case with all other extracts), the first five minutes of the movement were copied on to a master CD, followed by a 5 - 10 second logarithmic fade-out at a musically appropriate moment within the following 30 seconds.[7] Each recorded extract was finally transferred to the computer (a PC with a SoundBlaster 128 sound card) responsible for presenting stimuli to participants via Sony Digital MDR-V4 headphones.

A uniquely ordered "programme" of extracts was created for each experiment session by printing numbers corresponding to each extract in a random order; randomisation was carried out by a simple computer program written for the purpose. Before each experiment session, the "programme" was used to inform the computer of the order in which to play the extracts. Following this order, extracts were presented in succession with a 20 second gap between each.

[7]Both the Sloboda (1991) and Peretz et al. (1998) papers refer to specific bars within each musical work. However, as the present experiment required relatively long and coherent extracts, and most of the works in question are relatively consistent in their intensity and mood, it was felt that taking the first part of a movement was appropriate.

Stimuli: stories

Choice of short stories was determined by the following criteria: each should take the average reader approximately 5 minutes to read;[8] each should be sophisticated enough to be interesting to the reader but easily comprehensible; between them, they should be differentiated enough to elicit a range of responses in a typical reader; they should be varied in style. As there was no previous experimental work available to draw on when making the choices, the experimenter selected twelve stories which he felt matched the criteria. Although the potential for idiosyncratic decisions was clearly present, this was not considered to be a major problem, as the experiment results themselves would reveal whether the criteria had been met (as indeed was the case with the musical extracts).

The following short stories were chosen:[9]

1. Chandler - *Intaglio*
2. Chandler - *The Window*
3. Hemingway - *Cat in the Rain*
4. Hemingway - *The Old Man at the Bridge*
5. Lively - *Marriage Lines*
6. Lively - *The Dream Merchant*
7. O'Flaherty - *Three Lambs*
8. Henry - *A Newspaper Story*
9. Henry - *The Voice of the City*
10. Poe - *The Masque of the Red Death*
11. Poe - *The Tell-tale Heart*
12. Thomas - *The Dress*

[8] By necessity, this requirement was interpreted somewhat loosely. In any event, informal observation of participants reading the stories showed a very large difference between fast and slow readers.

[9] Full bibliographic information about the source of each story can be found in Appendix B.

It is worth noting that each of these stories, whilst clearly potentially moving, is somewhat ambiguous: none can be categorised as, say "happy" or "sad"; instead, as is the case with many short stories, each contains conflicting signals and can be read in many ways. Whilst this will complicate analysis of the second part of the experiment where music and stories are experienced simultaneously—it will not be possible unambiguously to determine the effect of context valence on the experience of music—it does increase ecological validity, as complexity and ambiguity are typical attributes of non-fictional situations and contexts.

Having been chosen, the stories were scanned into a computer with OmniPage optical character recognition software, checked for errors, then formatted using Palatino 12pt font. Each had its title and attribution removed and was assigned instead a label (a letter between "A" and "L") before being printed on sheets of A4 paper using a 1200dpi Laser Printer. In their reformatted form, the stories took between 2 and 5 sheets of paper; the pages of each story were stapled together in the top left-hand corner.

Before each experiment session, the stories were piled in a randomized sequence to guard against serious order effects; the randomisation was generated by the same computer program used to randomise the music ordering.

Participants

Eleven people participated in this part of the experiment. All were either students or staff at Cambridge University. None of the students was reading psychology; 3 studied music. Five participants were male; 6 were female. Ages ranged from 19 to 34 years, with a mean of 23.7 years and a standard deviation of 4.3 years.

Procedure

Participants carried out the experiment task individually, in a comfortable, quiet office-like environment. On arrival in the experiment room, each was assigned to one of two groups: Group A participants read the texts first, and then listened to the musical extracts; Group B participants listened to the musical extracts first and then read the texts. Although it was hoped that the order in which the tasks were performed would be unimportant, this grouping was implemented to guard against fatigue effects. 6 participants were allocated to Group A and 5 to group B.

Group A participants were asked to sit down and read through the following set of instructions:

Thank you for agreeing to participate in this experiment. These instructions will tell you everything you need to know about the tasks involved. Please read them carefully, taking as much time as you need. If you have any questions or need any clarification, please do not hesitate to ask the experimenter. The experiment is split into four sections, two of which are very brief, and two of which will take somewhat longer. In total, the experiment should take approximately 2 hours and 10 minutes, including a 5-minute break after Part Two.

Section One

On the table in front of you, you will see a form entitled "Response Form A". The task in this part of the experiment is simply to answer each of the questions on "Section One" of the form. Once you have done this, please inform the experimenter.

Section Two

You will be presented with a pile of twelve texts, each of which contains a complete short story. The task is to read each text and, after each one, answer the five simple questions you will find printed on a grid in "Section Two" of "Response Form A". Please read the five questions now and ask the experimenter if you would like any clarification.

Each text is clearly labelled with a letter between "A" and "L". The grid on the response sheet contains twelve rows, also labelled between "A" and "L". When answering a questions, please put the answer in the grid cell that corresponds to the column of the question you are answering, and the row whose label matches the label of the text you have just read. In other words, if the first text you read is "Text C", your answer to the second question would go in the cell which is in the third row down, and the second column across. Please indicate your answer by circling the relevant response option.

You may take as much time as you need to read the texts. Please inform the experimenter when you have finished. PLEASE BE SURE TO READ THE TEXTS IN THE ORDER IN WHICH THEY WERE GIVEN TO YOU; DO NOT SHUFFLE THEM FIRST!

Section Three

You will be presented with a series of twelve musical extracts with a short break between them. The task is to listen to each extract and, after each one, answer the five simple questions you will find printed on a grid in "Section Three" of "Response Form A". Please read each question now and ask the experimenter if you would like any clarification.

You will also be presented with a sheet of paper on which is printed a list of twelve items labelled "Extract 1" through "Extract 12". Think of this list as a programme: the items on it (and the order in which they occur) correspond to the musical extracts that you will hear (e.g. the third item in the list corresponds to the third extract that you hear). The grid on the response sheet contains twelve rows, also labelled between "1" and "12". When answering a questions, please put the answer in the grid cell that corresponds to the column of the question you are answering, and the row whose label corresponds to the extract you have just heard. In other words, if the first extract you hear is labelled "Extract 5", your answer to the second question would go in the cell which is in the fifth row down, and the second column across. Please indicate your answer by clearly circling the relevant response option.

If you would like a short break between any extracts, please tell the experimenter who will pause the music.

Section Four

Please answer the question printed in "Section Four" of "Response Form A", and inform the experimenter when you have finished.

The questions in Section One asked about participants' age, gender, and time spent regularly listening to music and reading; the Section Four question merely asked participants to write "any general comments concerning the texts, the extracts, the questions, your responses, or any aspect of the task you have been asked to perform".[10]

Before embarking on the experiment itself, the experimenter asked participants to confirm that they had understood what was required of them. In addition, due to a query raised by the first participant who completed the experiment, it was re-iterated that whereas an answer of "neither" to the second question in each group (the "valence" question) implied a

[10]The original instruction sheet and response forms actually referred to four distinct "parts" of the experiment rather than "sections". The word "part" has been substituted with "section" here to disambiguate descriptions, because the word "part" is used in this chapter to refer to the two separate halves of the experiment (music and stories presented separately vs. music and stories presented together).

low score in the first question (the "movingness" question), an answer of "don't know" did not exclusively imply either a low or a high score. Once participant and experimenter were happy, the former was invited to start reading the texts.

When participants announced that they were ready to start the listening part of the experiment, they were encouraged to sit comfortably, and asked to wear a pair of headphones (these were connected to the computer via a 7 metre lead, affording considerable freedom of movement). Once the experimenter was convinced that a participant had understood how to read the "programme", and would therefore write responses in the correct box, he programmed the computer with the randomisation pattern printed on the programme, and the extracts started to play (with a 20 second gap between each). Despite the allowance made in the instructions, no participants asked for the music to be paused at any time.

At the end of the session, participants were briefed on the purpose of the experiment, thanked for their time and were allowed to ask any questions or share any comments with the experimenter.

The protocol for participants assigned to Group B was identical to that for Group A with the exception that the second and third parts of the experiment were inverted so that these participants heard the musical extracts, and then read the texts.

Design

The design consisted of 6 between-subject variables and 2 separate within-subject variables; there were 10 dependent variables. Between-group variables were:

AGE: participant age (parametric: years).

GENDER: participant gender (categorical: male or female).

LIST: whether participant ever listens to music (categorical: yes or no).

HRSLIST: number of hours participant spends listening to music per week (parametric: hours).

READ: whether participant ever reads for pleasure (categorical: yes or no).

HRSREAD: number of hours participant spends reading for pleasure per week (parametric: hours).

The within-group variables were:

MUSIC: the musical extract being played (categorical: 12 categories).

TEXT: the text being read (categorical: 12 categories).

The dependent variables were:

MV_MUS: "movingness" rating for a musical extract. i.e. answer to the first question (parametric: 7pt scale).

VL_MUS: "valence" of a musical extract. This is derived from participants' answer to the second question and parameterized as follows: an answer of "positive" is encoded as a score of +1; "negative" is encoded as -1; "neither" and "don't know" are both encoded as 0. Thus, this value can be considered to be a parametric positive/negative "bias" rating.

EN_MUS: "enjoyment" rating for a musical extract. i.e. answer to the third question (parametric: 7pt scale).

QT_MUS: "quality" rating for a musical extract. i.e. answer to the fourth question (parametric: 7pt scale).

RB_MUS: whether a participant has heard a musical extract before (categorical: yes or no)

MV_TXT: "movingness" rating for a story. i.e. answer to the first question (parametric: 7pt scale).

VL_TXT: "valence" of a story. This is derived from participants' answer to the second question and parameterized as follows: an answer of "positive" is encoded as a score of +1; "negative" is encoded as -1; "neither" and "don't know" are both encoded as 0. Thus, this value can be considered to be a parametric positive/negative "bias" rating.

EN_TXT: "enjoyment" rating for a story. i.e. answer to the third question (parametric: 7pt scale).

QT_TXT: "quality" rating for a story. i.e. answer to the fourth question (parametric: 7pt scale).

RB_TXT: whether a participant has read a story before (categorical: yes or no).

Part One Results and Discussion

The following analysis is split into two main parts: the first part considers the relationship between MUSIC, TEXT and the between-subject variables (AGE, GENDER, LIST, HRSLIST, READ, HRSREAD) on the main dependents (MV_MUS, VL_MUS, MV_TXT and VL_TXT); it also examines relationships between the dependent variables themselves (MV_MUS, VL_MUS, EN_MUS, QT_MUS, RB_MUS, MV_TXT, VL_TXT, EN_TXT, QT_TXT and RB_TXT). The second part analyses MUSIC and TEXT, MV_MUS, VL_MUS, MV_TXT and VL_TXT in more detail, with the aim of categorising the stimuli for Part Two of the experiment.[11]

Preliminary examination of the parametric dependent variables immediately revealed that all exhibited similar distributions (with the exception of VL_MUS and VL_TXT, whose scale is somewhat different):

Variable name	Mean Value	Std. Deviation	Minimum	Maximum
MV_MUS (music movingness)	3.90	1.74	1	7
VL_MUS (music valence)	-0.13	0.82	-1	1
EN_MUS (music enjoyment)	3.92	1.53	1	7
GS_MUS (music quality)	3.99	1.43	1	7
MV_TXT (story movingness)	4.36	1.39	1	7
VL_TXT (story valence)	0.18	0.85	-1	1
EN_TXT (story enjoyment)	4.58	1.36	1	7
GS_TXT (story quality)	4.64	1.34	1	7

General analyses: musical extracts

The first set of analyses examines the effect of GROUP, SEX and MUSIC on MV_MUS. A three-way Analysis of Variance (two-between; one-within; type III sums of squares) showed no significant effects of GROUP or SEX, but a significant effect of MUSIC: $F_{(11,77)} = 2.08$; $p < 0.031$. Reassuringly, then, the different musical extracts were rated significantly differently in terms of movingness. There were no significant two- or three-way interactions.

[11] All analyses were performed using *Statistica* software from Statsoft Inc.

A three-way ANOVA (two-between; one-within; type III sums of squares) of GROUP, SEX and MUSIC on VL_MUS showed no significant effect of GROUP, but a significant effect of SEX: $F_{(1,6)} = 18.35$; $p < 0.05$ and an extremely highly significant effect of MUSIC: $F_{(11,66)} = 8.79$; $p < 0.0000001$.[12] Further examination of the effect of SEX revealed that the mean VL_MUS score for female participants was 0.41 compared with 0.01 for male participants; in other words, female participants tended to be more positively moved than male participants by the musical extracts. Although this statistic presents no obvious explanation, one possibility suggested by informal observation is that female participants tended to view the experiment task itself in a more positive light. The extremely highly significant effect of MUSIC indicates that participants were easily able to differentiate between the different extracts in terms of their valence. There were no significant two- or three-way interactions.

All participants reported that they listened to music regularly, obviating the need for an analysis in terms of the categorical variable, LIST: the minimum number of hours spent listening was 1.5 hours per week; one participant listened for as many as 35 hours; the mean amount was 12.3 hours with a standard deviation of 9.5.[13] Pearson's Product-Moment Correlation revealed no significant correlation between HRSLIST and mean MV_MUS (that is, the average MV_MUS score across participants for a given extract), or between HRSLIST and mean VL_MUS (average VL_MUS score across participants for a given extract). Neither was there any significant correlation revealed between AGE and MV_MUS, or between AGE and VL_MUS.

Analysis of the overall relationship between MV_MUS and EN_MUS, and between MV_MUS and GS_MUS was carried out by concatenating scores from all participants and all extracts into a single 132-case column of datapoints for each variable, [14] and performing a Pearson's Product-Moment Correlation on the resulting dataset. This procedure revealed a significant correlation between MV_MUS and EN_MUS ($r = 0.70$; $p < 0.05$) and between MV_MUS and GS_MUS ($r = 0.63$; $p < 0.05$). In other words, musical extracts which were found to be moving were also found to be enjoyable and were considered to be good music.[15]

[12] One participant omitted to give a value for VL_MUS for one of the extracts; his answers were therefore removed from this analysis.

[13] Participants who regularly listened to music whilst carrying out other tasks were told to estimate the number of hours spent "background listening", and divide this number by two in order to work out a total weekly listening figure.

[14] For each variable: 12 extracts x 11 participants = 132 cases.

[15] A scatterplot between MV_MUS and EN_MUS revealed revealed the correlation to be surprisingly linear, and not at all suggestive of the Berlyne (1971) inverted-U. A possible reason for this finding is that the experience

As a Cochran's Q Test on RB_MUS revealed highly significant differences between extracts in terms of how well they were known: $Q_{(df=11)} = 44.56$; $p < 0.000006$, it seemed appropriate to analyse the effect of RB_MUS on MV_MUS and VL_MUS. Each extract, therefore, was given a score corresponding to the percentage of partipants who answered "yes" to RB_MUS. These scores were subjected to a Pearson's Product-Moment correlation test against each extract's mean MV_MUS score, yielding no significant results. However, analysis of the transformed RB_MUS score with VL_MUS did reveal a significant correlation: $r = 0.72$; $p < 0.05$. Further examination revealed this to be the result of polarized clusters of datapoints: participants knew the negatively valenced extracts least well and positively valenced extracts better; or perhaps familiarity with an extract led to it being perceived as more positively valenced![16]

General analyses : short stories

A three-way Analysis of Variance (two-between; one-within; type III sums of squares) revealed that neither GROUP nor SEX significantly affected the degree to which participants were moved when reading short stories (MV_TXT). Just as with the musical extracts, participants did differentiate between the stories, as indicated by an extremely highly significant effect of TEXT: $F_{(11,77)} = 6.56$; $p < 0.0000001$. There were no significant two- or three-way interactions.

A three-way ANOVA (two-between; one-within; type III sums of squares) of GROUP, SEX and TEXT on VL_TXT showed no significant effect for GROUP, but a marginally significant effect of SEX: $F_{(1,7)} = 3.65$; $p < 0.098$; it appears that female participants tended to rate the stories as more clearly negatively valenced than did male participants (mean VL_MUS score of -0.28 for female participants as opposed to -0.028 for male participants). Although it may be possible to explain this discrepancy in terms of gender-biased interpretations of the texts, the marginal significance of the effect hardly warrants such treatment. As expected, TEXT had an extremely highly significant effect on VL_TXT: $F_{(11,77)} = 7.60$; $p < 0.0000001$.

of listening to these musical extracts in the experimental environment did not evoke strong enough arousal for inverted-U trends to emerge.

[16]There is not enough data here unambiguously to support either hypothesis. However, the latter explanation seems plausible: it is an oft-cited folklore amongst musicians that the experience of listening to a stereotypically negatively valenced but familiar piece of music is often a highly positive one. A significant correlation between the transformed RB_MUS scores and mean EN_MUS (Pearson's Product-Moment: $r = 0.58$; $p < 0.05$) supports this observation further.

This effect, together with the equivalent effect of TEXT on VM_TXT, will be discussed further below.

All participants reported that they read regularly for pleasure, so there was no need to analyse the effect of the variable READ; however, the number of hours per week participants tended to spend reading for pleasure was markedly lower than the amount of time spent listening to music, with a minimum of 1 hour, a maximum of 14 hours, a mean of 6.3 hours with a standard deviation of 4.1 hours. Pearson's Product-Moment showed no significant correlation between HRSREAD and MV_TXT, or between HRSREAD and VL_TXT. Participant age was also correlated neither with MV_TXT nor with VL_TXT.

An analysis of the overall relationship between MV_TXT and EN_TXT, and between MV_TXT and GS_TXT was carried out using the same procedure as the equivalent analysis with musical extracts: single 132-case columns of datapoints were generated for each variable by concatenating scores from all participants and all extracts. MV_TXT and EN_TXT were highly correlated: $r = 0.68$; $p < 0.05$; MV_TXT and GS_TXT were very highly correlated: $r = 0.80$; $p < 0.05$. In short, participants were moved by stories that were enjoyable and found to be good; or perhaps more plausibly, enjoyable, good stories were the most moving. That the correlations seen here are notably stronger than their equivalents relating to the musical extracts can probably be attributed to the relative conceptual explicitness of a short story compared with a musical extract.[17]

None of the participants reported having read any of the short stories previous to the experiment, obviating the need for any analysis involving RB_TXT.

Categorisation of Musical Extracts

This part of the analysis considers observable differences between the various musical extracts, based on MV_MUS and VL_MUS ratings; its aim is to isolate patterns where they exist in the data and to characterise and categorise the extracts according to participants' emotional responses. As we are interested solely in differences that can be attributed to the effect of MUSIC, and cannot call on the services of an ANOVA to factor out variability due to participants' use of the scales, all data presented here was normalized (converted to Z-scores) prior to further analysis.

[17]Just as with the musical counterpart of this analysis, no inverted-U trends were seen in these data.

The two variables MV_MUS and VL_MUS define between them a two-dimensional "emotion-space" within which each musical extract can be said to be located. In order to examine this space, mean MV_MUS and VL_MUS scores were derived for each extract—by taking the average score across participants—and plotted on a scattergraph (see figure 7.1). Two distinct clusters of extracts are immediately evident:

- One cluster, containing the Beethoven Symphony, the Handel and the Mozart, appears at the top right-hand corner of the graph, indicating a high mean MV_MUS and VL_MUS score. In other words, participants, on average, found these extracts highly moving and clearly positively valenced. These ratings fit well with prior conceptions: all three of these extracts are fast-moving, in the major key, and have been generally found by listeners to be moving and happy, according to prior experimental work (see page 117 above).

- Another cluster, containing the Rachmaninov and Beethoven concertos, and the Schoenberg, appears at the bottom right-hand corner of the graph, indicating a high mean MV_MUS and a low mean VL_MUS score. In other words, participants, on average, found these extracts to be highly moving and clearly negatively valenced. Again, these ratings fit with prior conceptions: all three extracts are slow-moving, in the minor key and, according to prior experimental work, have been found to be sad, or moving, or both (see page 117 above).

The remainder of the extracts do not fall so obviously into categories, although some general patterns are worthy of mention. First, extracts whose VL_MUS score approached zero did not score high MV_MUS values either; i.e. extracts whose valence was unclear were never highly moving.[18] Second, with a single exception (the Shostakovich), extracts whose mean VL_MUS scores were far from zero always reached at least an average MV_MUS score; i.e. music which had clear valence was never totally unmoving.

All three of the extracts that moved participants but with no particular valence (that is, the Brahms, the Rachmaninov Symphony, and the Sibelius) had been reported as highly moving pieces in Sloboda (1991), and for that matter by the present experimenter, whereas

[18] This result seems to run somewhat contrary to the folk-musicological observation listeners can be highly moved by music but undecided as to valence. Instead, the results here seem to suggest that whereas moderately moving music can be undifferentiated with respect to valence, being moved to a greater extent involves disambiguation of valence.

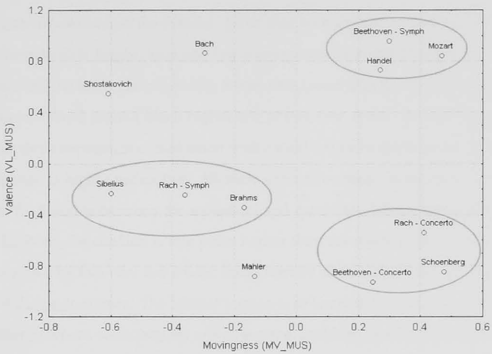

Figure 7.1: MV_MUS against VL_MUS - plots of normalized means

they were only judged as moderately moving by participants. A possible explanation for the difference seen here is that the respondants to Sloboda's questionnaire were mostly highly-trained musicians, for whom the pieces would have been very familiar. Although no significant correlation was seen between MV_MUS and RB_MUS in the present experiment, it is possible that these pieces (all of which were judged as relatively unfamiliar) would have appeared further toward to right-hand side of the graph had more of the partipants been familiar with them.

Three extracts appear to be lone outliers, namely, the Bach, Mahler and Shostakovich extracts. The position of these extracts highlights an interesting asymmetry: first, whilst highly moving music was always clearly valenced, it was not the case that clearly valenced music was always highly moving. A possible explanation is that in cases where participants were not particularly moved by a musical extract, valence scores reflected awareness of a composer's likely intentions (ie. reflected recognition of a musical style). Given the small number of extracts involved, however, a further experiment would be required to explore that hypothesis further.

Considered in the light of the actual musical content, all three of these outliers fall in somewhat unexpected places in the "emotion-space": it was expected that the Bach extract, found here to be moderately moving but clearly valenced, should score similar ratings along both dimensions as the Handel, given that both are similar in style, tempo and mode; the Shostakovich, bright, loud and fast was expected to evoke a high MV_MUS score, but in fact evoked the lowest of all; finally, it was anticipated that the Mahler, complex, minor and intense, would score a much higher MV_SCORE than it did. Examination of the distributions for these extracts, in combination with consideration of informal comments made by participants, is highly instructive. All three extracts evoked a wide range of scores (a difference of 5 - 6 points between the minimum and maximum before normalisation). In the case of the Bach, the median is one point higher than the mean (5 as opposed to 4); the mean is lowered by two very low scores, both given by participants who commented on their dislike of the performance. The Mahler's mean was lowered by a second pair of participants, neither of whom were familiar with the piece, and both of whom found it tedious. As for the Shostakovich, which evoked a wide range of MV_MUS ratings but a relatively even spread, several participants informed the experimenter that they had found it unmoving because of its triteness!

Given this description of scores and informal evidence in accounting for the unexpected positions of the three outlying extracts, it could be argued that they are not outliers at all. They are certainly not anomalies in any real sense; rather, they illustrate the importance of context (in the form of idiosyncratic listener effects) on the ability of a musical work to evoke an emotional response. However, as the main purpose of this part of the experiment was to attempt to find "clusters" of musical extracts that could be used interchangeably in the second part, these strays are unhelpful. Removing them leaves three distinct clusters of extracts that could potentially be used in the second part:

1. Highly moving; positively valenced
   — Beethoven Symphony, Handel, Mozart
2. Highly moving; negatively valenced
   — Beethoven Concerto, Rachmaninov Concerto, Schoenberg
3. Moderatively moving; non-valenced
   — Brahms, Rachmaninov Symphony, Sibelius

In order to test the robustness of this otherwise *ad hoc* clustering, a K-Means Cluster Analysis was performed. First, normalized scores for MV_MUS and VL_MUS were concatenated, producing 22 data points (11 participants x 2 variables) to act as dimensions for each extract. These data were then subjected to K-Means clustering, using 3 means and starting with maximized between-cluster distances; the Bach, Handel and Shostakovich extracts (discussed above) were not included in the clustering. After 2 iterations, 3 clusters were found, whose contents matched the *ad hoc* clusters. The smallest Euclidean distance between clusters was 0.96 between clusters 2 and 3. In short, the results of this analysis support the clustering of nine musical extracts into three distinct groups, ready for use in the second part of the experiment.

Categorisation of Short Stories

This final part of the analysis considers observable differences between the various short stories, based on MV_TXT and VL_TXT ratings; its aim is to isolate patterns where they exist in the data and to characterise and categorise the stories according to participants' emotional responses. As we are interested solely in differences that can be attributed to the effect of TEXT, and cannot call on the services of an ANOVA to factor out variability due to partipants' use of the scales, all data presented presented here were normalized (converted to Z-scores) prior to further analysis.

The analysis here parallels identically that of the musical extracts: the two variables MV_TXT and VL_TXT define between them a two-dimensional "emotion space" within which each short story is situated; this space was examined by deriving means for MV_MUS and VL_MUS scores for each extract, and plotting the results on a scattergraph (see figure 7.2). The pattern of means is quite different to the musical equivalent; notably, with two exceptions (*The Cat in the Rain* and *Intaglio*) none of the stories lies near the zero-valence line, indicating that even if participants were hardly moved at all by a story, valence was always unambiguous.[19] Another major difference is that none of the stories fall in the top right-hand quadrant of the graph; in other words none of them were found to be highly moving and positively valenced. Whether these distinctions are due to inherent differences between listening processes and reading processes, or whether they are merely a function of the specific

[19]The result could also be the effect of very wide disagreement amongst the subjects, but examination of distributions showed that this was not the case; in fact, the only two extracts that evoked a wide spread of responses were *The Cat in the Rain* and *Intaglio* with standard deviations (non-normalized) of 0.83 on a mean of -0.09 and 0.92 on a mean of -0.36 respectively.

stories being used in this experiment cannot, sadly, be resolved without a much larger-scale experiment.

The stories appear to be grouped into four distinct clusters, which can be summarized as follows:

- One set of stories, containing *Intaglio*, *The Window*, and *The Tell-tale Heart*, resides at the bottom right-hand corner of the graph. This set is characterised by high MV_TXT scores and negative VL_TXT scores; in other words, participants on average found these stories moving and negatively valenced.
- A second set of stories, including *The Dress*, *The Masque of the Red Death*, and *The Old Man at the Bridge*, were characterised as moderately moving and clearly negative (moderate MV_TXT scores and clearly negative VL_TXT scores).
- *The Dream Merchant*, *Marriage Lines*, and *The Three Lambs*, were characterised as moderately moving and positively valenced.
- A final cluster, containing *A Newspaper Story*, *The Voice of the City*, and *The Cat in the Rain*, were found to be not particularly moving and positively valenced.

The distribution of stories across "emotion space" is comfortably wide. In addition, there appear to be no particular anomalies or outliers; all stories appear in unremarkable locations as far as the experimenter was concerned. However, unlike the musical extracts, emotional response to these stories has not, to the best of the experimenter's knowledge, been the subject of previous psychological investigation. Any attempt, therefore, to compare the spread of datapoints here with any prior conceptions must remain purely idiosyncratic.

Despite the apparent differentiation between clusters on the scattergraph, a K-Means Cluster Analysis (executed using exactly the same procedure as that for the musical extracts, but this time searching for 4 clusters instead of 3) produced surprising and rather unsatisfactory results! One cluster contained five stories (all whose means indicated positive VL_TXT scores), one contained *The Tell-Tale Heart* and *The Window* (two of stories in the bottom, right-hand side of the graph, notably excluding *Intaglio*), the third contained the moderately moving, negatively valenced stories (*The Dress*, *The Masque of the Red Death*, and *The Old Man at the Bridge*) plus *Intaglio* (which effectively migrated from its more logical place in the bottom

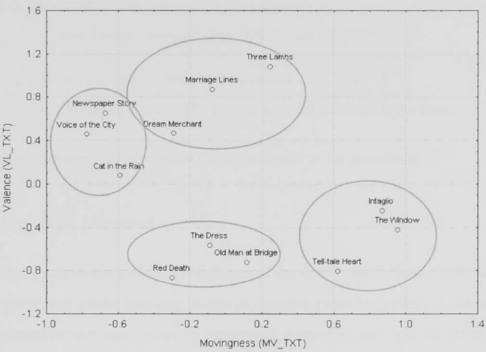

Figure 7.2: MV_TXT against VL_TXT - plots of means

right-hand corner), leaving *The Cat in the Rain* as the sole member of the fourth cluster. In short, *Intaglio* and *The Cat in the Rain* were "misplaced", and the analysis did not differentiate along the MV_TXT dimension between low and medium scores.

It was suspected that this counter-intuitive clustering was caused by participants' disparate responses to the valence of the two "misplaced" stories (indicated by high VL_MUS variance; see page 133) affecting the clustering algorithm. The K-Means Cluster Analysis was therefore re-run without *Intaglio* and *The Cat in the Rain*. This time, the clusters turned out exactly as expected: one cluster contained the stories judged to be negatively valenced and moving, one those which were negatively valenced and moderately moving, one those which were positively valenced and moderately moving, and one those which were positive and not very moving. The minimum Euclidean distance between two clusters was 0.93.

In conclusion, the second K-Means Cluster Analysis supported the *ad hoc* clustering of short stories into 4 clusters, based on partipants' responses:

1. Highly moving; negatively valenced
   — *The Window, Tell-tale Heart,* (*Intaglio*)

2. Moderately moving; negatively valenced
   — *The Dress, The Old Man at the Bridge, The Masque of the Red Death*

3. Moderately moving; positively valenced

   — *Three Lambs, Marriage Lines, The Dream Merchant*

4. Not very moving; positively valenced

   — *A Newspaper Story, The Voice of the City, (The Cat in the Rain)*

These clusters will be used in the second part of the experiment.[20]

Part Two Method

Participants were invited to read each of the twelve short stories used in the first part of the experiment whilst listening to one of the nine remaining musical extracts (for each story, a musical extract was chosen by computer). After completing each music / story pair, participants were asked to answer each of the following questions:

1. *How moving did you find the music (on a scale of 1 to 7, where 1 = as unmoving as any that I have ever heard; 7 = as moving as any that I have ever heard)?*

2. *If you found the music moving, was it positively moving, negatively moving or both / neither (positive = e.g. Made me happy, amused, etc...; negative = e.g. made me sad, morose, contemplative...; both / neither = It was not unambiguously positive or negative; don't know)?*

3. *Have you heard this piece of music before?*

4. *How moving did you find the story (on a scale of 1 to 7, where 1 = as unmoving as any that I have ever read; 7 = as moving as any that I have ever read)?*

5. *If you found the story moving, was it positively moving, negatively moving or both / neither (positive = e.g. Made me happy, amused, etc...; negative = e.g. made me sad, morose, contemplative...; both / neither = It was not unambiguously positive or negative; don't know)?*

6. *Have you heard this story before?*

[20]Despite the problematic nature of VL_TXT scores associated with *Intaglio* and *The Cat in the Rain*, it was decided not to exclude them from the clusters which would be used in the next part of the experiment for two reasons: first, it would be interesting to see whether the disagreement could be resolved by a second set of participants; second, if they continue to be mathematically, if not visually anomalous, it would be trivial to exclude these extracts from analysis later on.

7. *Did you feel that reading the story and listening to the music were very separate experiences (i.e. your emotional response to the text was separate from your response to the music), or did you respond to music and text as an integrated whole? (completely separate, slightly separate, fairly integrated, completely integrated)?*

A complete experiment session lasted approximately 1 hour and 20 minutes, depending on the reading speed of the participants.

The Questions

In contrast to the first part of the experiment, the range of questions asked here about the musical extracts and stories was restricted to the two that asked specifically about participants' emotional responses to the stimuli (1,2 and 4,5) and to a simple dichotomous question asking about familiarity of the musical extracts and stories (3,6). The rationale for excluding both questions asking about participants' enjoyment of the stories and those asking for objective assessments of the stimuli was two-fold: first, if participants were asked all the questions that appeared in the first part of the experiment, they would have to answer a total of eleven questions per trial, including six rating scales; this was felt to be a slightly unrealistic requirement, likely to lead to arbitrary or indiscriminate use of the scales. Second, analysis of the first part of the experiment provided persuasive evidence that measures of "movingness", enjoyment and objective quality provide highly redundant information and are therefore unnecessary (see page 126).

Part of the purpose of the extra questions in the first part of the experiment had been to provide a relief against which to ensure that participants fully understood the nature of the "movingness" and "valence" questions, namely, that they asked about personal reactions to the music and stories rather than assessment of, say, a writer's or composer's intention (see page 117). In order to compensate for their lack here, the following instructions were given to participants when they read the questions for the first time:

> *When replying to questions 1 and 4, please note that you are not being asked to make an objective judgement on whether you think the text / music intends to be moving; you are being asked specifically about your personal emotional reactions.*

In order to anticipate queries that were likely to emerge, participants were also told:

> *Do not worry if you find it difficult to separate your responses to the text and the music; it is perfectly acceptable to give the same ratings to both when answering the questions. Likewise, do not worry if your gut reaction seems to be at odds with what you might objectively have expected it to be: an honest report of your response to the reading / listening experience is far more useful than a report of what you think it should have been!*

It was anticipated that some participants would find the experiences of listening to music and reading a story quite separate, whereas for others they would be integrated; the final question in each set was intended to give the experimenter some insight into this. A four-way forced-choice question was used here in place of the more traditional rating scale because discussion with potential (but not actual) participants revealed that integratedness of music and text was a difficult thing to assess; it was felt that a numerical 7-point scale might lead to arbitrary results.

Stimuli

Stimuli consisted of short stories coupled with musical extracts: the stories were those used in Part One of the experiment, presented in an identical format; the musical extracts were the nine identified in the Part One categorisation analysis (see page 129), also presented in identical format with the exception that sound playback was here controlled by different software (see page 140 below).

A trial was produced by pairing music extracts with stories according to the following procedure:

1. Music extracts were assigned to one of three groups, according to the 3-way categorisation defined by the Part One music categorisation analysis (see page 129).

2. Stories were assigned to one of four groups, according to the 4-way categorisation defined by the Part One story analysis (see page 133).

3. For every experiment session, the order of the music extracts and stories was randomised *within each group*, so that the groups contained the same members but in a different permutation for each participant.

4. For each of the 12 possible combination of music extract groups and short story groups (e.g. music group A with story group A, music group A with story group B, etc...), a representative music extract and text were chosen by cycling through the members within each group and using successive members every time that group was involved in a pairing. To clarify, if the music extract group were labelled A, B and C, and the story groups were labelled I, II, III and IV, and the music / stories within each group were labelled i, ii, and iii, extract pairings would have been as follows: Ai with Ii, Aii with IIi, Aiii with IIIi, Ai with IVi, Bi with Iii, Bi with IIii, ...etc.).

5. For every experiment session, the resultant list of music / story pairings was then itself randomised.

The effect of this procedure was three-fold: first, it ensured that every participant read every story once, and heard every music extract either once or (in three cases) twice; second, the pairings of specific musical extracts with specific texts would vary from participant to participant; third, every participant read an example of each of the categories of story whilst hearing an example from each of the categories of music. Each participant heard three extracts twice by necessity, as there were three fewer extracts than stories; this was felt to be non-problematic, as the likely familiarity of many of the extracts for many participants, coupled with the non-conceptual nature of music as opposed to text, should ameliorate any effects of rehearing.

Presentation of the stimuli was handled by a Windows PC running *The MuPsych Lab* software;[21] sound was presented through the same hardware as in Part One of the experiment; the stories were presented in paper form, as it was felt that participants would not enjoy reading from a screen; the software displayed a legend detailing which story should be read for each trial.

Music extracts all lasted approximately five minutes (see page 117), yet the time taken to read each story must vary from participant to participant. It was highly likely that several participants would take longer to read a story than the length of the corresponding music extract; this unavoidable situation was beyond the realms of experimental control. In order to provide symmetry for any potential effects, it was decided to allow the converse: in trials

[21]See Appendix A for a brief description of *The MuPsych Lab* software, and a listing of the control code that was used to run this experiment.

where music extracts outlasted the length of time taken to read a story, the music continued to play until the extract ended, even though curtailing this would have been within the bounds of experimental control.

Participants

Twelve people participated in the experiment. All were either students or staff at Cambridge University. None of the students was reading psychology; 2 studied music. Seven participants were male; 5 were female. Ages ranged from 19 to 56 years, with a mean of 30.6 years and a standard deviation of 13.2 years. Noone who participated in this part of the experiment had been a participant in Part One.

Procedure

People participated in the experiment individually; all sessions took place in the same room as Part One; participants sat in the same chair at the same desk. On arrival, each participant was assigned to one of two groups ("Group A" or "Group B"). Protocol was identical for the two groups except for the order of the questions presented after each trial: Group A participants were first asked to report their responses to the music extracts, and then to the stories—i.e. questions were presented in the order printed above (see page 136)—whereas Group B participants were asked to record their responses to the stories first, and then to the music extracts.

All aspects of each experiment session were controlled by a Windows PC running *The MuPsych Lab* software; the software was responsible not only for playing the musical extracts and telling the participants the order in which the stories should be read, but also for presenting the questions after each trial and collecting responses.

A session started by displaying a set of instructions explaining the task and introducing the questions that would have to be answered after each trial; at this point, participants were invited to ask the experimenter if there was any uncertainty surrounding any aspect of the task, after which the following procedural instructions were presented:

> *On the desk next to you, you will see a pile of short stories, each of which is clearly labelled with a letter from "A" to "L"; the texts are stacked in alphabetical order. At the start of each trial, the computer will display the word "Text" followed by one of these letters; look*

> *through the pile of texts, and find the story labelled with the letter being displayed by the computer.*
>
> *Once you have found the text, use the computer mouse to click the button labelled "start trial". At this stage, a musical extract will start to play and you should start reading the story. You should read the story through to the end, even if the music has stopped by the time you have finished reading.*
>
> *As soon as you have finished reading the story, click the button marked "I have finished reading". If the music has not already finished, please wait patiently until it does.*
>
> *Once you have finished reading the story and the music has stopped playing, the computer will present you with each of the questions in turn. Answer each one by clicking on the appropriate selection and then clicking "OK". Once you have answered all the questions, the computer will ask you to click when you are ready for the next trial.*
>
> *Please ask the experimenter if you have any questions. Otherwise, you may put on the headphones that you will find on the desk in front of you, and begin in your own time.*

Once participants had confirmed that they had understood all of the instructions, they were invited to fill in an on-screen questionnaire asking exactly the same questions as in Section One of Part One of the experiment (see page 121). Once participants confirmed that they were wearing the headphones, *The MuPsych Lab* generated a unique trial ordering for the session (see page 138) and each trial was presented in succession, as per the instructions. After all had been completed, participants were thanked for participating and were invited to discuss any aspect of the experiment with the experimenter.

Design

The main design consisted of four between-group and two within-group variables. Between group variables were:

GROUP whether participants were assigned to Group A or Group B (categorical: A or B).

GENDER whether participants were male or female (categorical: m of f).

LIST: whether participant ever listens to music (categorical: yes or no).

READ: whether participant ever reads for pleasure (categorical: yes or no).

The within group variables were:

M_TYPE for each trial, this variable defined which category of music extract was used. Three possible conditions were: highly moving/positive, highly moving/negative, moderately moving/neither.

T_TYPE for each trial, this defined which category of story was used. Four possible conditions were: highly moving/negative, moderately moving/negative, moderately moving/positive, not moving/positive.

Dependent variables were:

MV_MUS: "movingness" rating for a musical extract. i.e. answer to the first question (parametric: 7pt scale).

VL_MUS: "valence" of a musical extract. This is derived from participants' answer to the second question and parameterized as follows: an answer of "positive" is encoded as a score of +1; "negative" is encoded as -1; "neither" and "don't know" are both encoded as 0. Thus, this value can be considered to be a parametric positive/negative "bias" rating.

RB_MUS: whether a participant has heard a musical extract before (categorical: yes or no)

MV_TXT: "movingness" rating for a story. i.e. answer to the first question (parametric: 7pt scale).

VL_TXT: "valence" of a story. This is derived from participants' answer to the second question and parameterized as follows: an answer of "positive" is encoded as a score of +1; "negative" is encoded as -1; "neither" and "don't know" are both encoded as 0. Thus, this value can be considered to be a parametric positive/negative "bias" rating.

RB_TXT: whether a participant has read a story before (categorical: yes or no).

SEP how separate the experience of reading and listening seemed. Four categories (completely separated, slightly separated, fairly integrated, completely integrated) were interpreted parametrically on a scale of 1 to 4, where high values indicated high separation.

COMPL the order in which participants completed their interaction with music extracts and story within a trial (categorical: music first, story first, together). If music extracts stopped within 15 seconds of a participant finishing to read, COMPL was interpreted as being "together".

REHEARD whether the music extract involved in a trial had been presented previously in the experiment (categorical: yes or no).

In addition, there were three potential covariates:

AGE: participant age (parametric: years).

HRSLIST: number of hours participant spends listening to music per week (parametric: hours).

HRSREAD: number of hours participant spends reading for pleasure per week (parametric: hours).

Finally, two variables—MUSIC (categorical: 9 conditions) and TEXT (categorical: 12 conditions)—referred to the individual music extract or text respectively that was presented in a specific trial. These two variables provide a means to examine relationships between music extracts and stories that does not depend solely on the categorisations derived in Part One of the experiment.

Part Two Results and Discussion

The analysis will start by examining the effect of M_TYPE, T_TYPE, GROUP and GENDER on MV_MUS, VL_MUS, MV_TXT, VL_TXT and SEP, using a series of Analyses of Variance. It will then attempt a more sensitive examination of the relationship between movingness and valence for music and stories, involving the data from Part One of the experiment.

Descriptive statistics for the dependent variables are presented in the table below. Overall, the responses were approximately normally distributed, covering the entire range of each scale. Of particular note, however, is the the high mean and small standard deviation of SEP; we will return to the implications of this later.

Variable Name	Mean Value	Std.Deviation	Minimum	Maximum
MV_MUS (music movingness)	4.21	1.32	1	7
VL_MUS (music valence)	0.12	0.80	-1	1
MV_TXT (story movingness)	3.92	1.50	1	7
VL_TXT (story valence)	-0.035	0.82	-1	1
SEP (separation between domains)	2.99	0.95	1	4

Analyses of Variance

A three-way repeated measures Analysis of Variance (two-between, one-within, type III sums of squares) revealed no significant effect of M_TYPE, T_TYPE, GROUP or GENDER on MV_MUS; there was, however, a marginal effect of M_TYPE ($F_{(2,16)} = 3.35$; $p < 0.06$). The surprising weakness of the M_TYPE effect probably reflects the nature of the music categorization process which by necessity clustered musical extracts with comparatively wide movingness ratings (see page 129).

In addition, there was a marginal two-way interaction between M_TYPE and GROUP ($F_{(2,16)} = 3.57$; $p < 0.052$), namely that participants who answered questions about the music extracts prior to answering questions above the text gave higher MV_MUS ratings for music in the "high/negative" category than participants who answered questions about the stories first. A tentative explanation might be that the movingness of high/negative music lost its potency once participants had been required to think about rating stories.

There was no two-way interaction between M_TYPE and T_TYPE.

Analysis of VL_MUS tells a rather different story: a three-way repeated measures ANOVA (two-between, one-within, type III sums of squares) revealed an extremely highly significant effect of M_TYPE ($F_{(2,16)} = 18.69$; $p < 0.000065$) and a significant effect of T_TYPE ($F_{(3,24)} = 3.46$; $p < 0.032$). The profile of the M_TYPE effect is exactly as might be expected: music in the "high/positive" evoked a high VL_MUS score, that in the "high/negative" category evoked a low score, and that in the "moderate/neither" category evoked a score that fell between the two. By contrast, the T_TYPE effect is of considerable interest: regardless of music type, participants gave lower VL_MUS ratings for music extract heard whilst reading stories in the "high/negative" category. In other words, powerful, negatively valenced stories were able to make participants' responses to the music seem more negative.

A significant three-way interaction involving M_TYPE, T_TYPE and GROUP ($F_{(6,48)} = 2.72$; $p < 0.023$) resists obvious explanation: it appears that in trials involving stories in the "moderate/negative" category, participants who answered questions about the stories first gave high VL_MUS scores for "high/positive" music and low scores for "high/negative" music compared with participants who answered questions about the music first. i.e. participants who answered questions about the stories first appeared to be more sensitive to differences between music types than the other participants.

A three-way repeated measures ANOVA (two-between, one-within, type III sums of squares) on MV_TXT revealed an extremely highly significant effect of T_TYPE ($F_{(3,24)} = 12.41$; $p < 0.00042$) and a marginal two-way interaction between M_TYPE and GROUP: $F_{(2,16)} = 3.53$; $p < 0.054$. The profile of the main effect was broadly as expected: the highest ratings were scored by "high/negative" stories, followed by "moderate/positive" ones, followed by "moderate/negative", followed by "not/positive"; post-hoc t-tests on group means revealed no significant differences between "moderate/positive" and "moderate/negative" groups.

The marginal two-way interaction between M_TYPE and GROUP was a result of participants who answered questions about the stories first giving significantly lower scores in trials containing music in the "moderate/neither" category than participants who answered questions about the music first. It seems plausible that regardless of the story type, participants were influenced insidiously by the movingness of the music when answering questions about the stories; however, participants who answered questions on the music first were able consciously to compensate.

Another three-way repeated measures ANOVA (two-between, one-within, type III sums of squares) on VL_TXT showed a highly significant effect of T_TYPE ($F_{(3,24)} = 22.33$; $p < 0.00001$) whose profile looks exactly as expected, simply following the mean valence of the categorization groups. A marginal two-way interaction between M_TYPE and T_TYPE ($F_{(6,48)} = 2.16$; $p < 0.064$) turned out to be rather less interesting than had been hoped: a series of post-hoc t-tests (rejecting at $p < 0.01$ to avoid Type I errors) revealed the significance merely to be the result of the difference between levels of T_TYPE not being significant at all levels of M_TYPE; no clear trends were suggested.

The final dependent to be subjected to a three-way repeated measures ANOVA (two-between, one-within, type III sums of squares) was SEP. A significant main of effect of M_TYPE ($F_{(2,16)} = 6.09$; $p < 0.011$) was the result of music in the "high/positive" category leading to significantly greater separation than music in the other groups. The implication is that people are more likely to fuse negative or ambiguous music than positive music with a story; another interpretation might be that people are more willing to accept a negative "commentary" on a text than a positive one. Such a finding is not particularly surprising, and could well be partly the result of habituation: the most common experiences in everyday life involving music and explicit narratives are those involving film, where moving negative or ambiguous music seems to form the mainstay of Hollywood tradition!

A three-way interaction between M_TYPE, T_TYPE and GENDER ($F_{(6,48)} = 2.89$; $p < 0.017$) defies obvious explanation; unfortunately, it is not possible to do post-hoc pairwise comparisons because of the lack of an unambiguous error term when between- and within-group variables are involved in an analysis.[22] However, the effect can be characterised as follows: for trials involving stories in the "high/negative" category, "high/positive" music evoked lower SEP scores for female participants than for male ones; conversely, "high/negative" music evoked lower SEP scores for male participants. For all other story types, female participants had a tendency to give lower SEP scores. Further investigation would be required to determine an accurate trend, but for the present it is appropriate merely to note that male and female participants appear to differ in the combinations of music and story type that are most likely to be fused into a single coherent experience.

Although the above ANOVA analyses yield some interesting findings, they are for the most part weak and difficult to interpret, a situation that is due largely to two inherent problems with the independent variables M_TYPE and T_TYPE. Whilst the music and story categorisations proved invaluable in ensuring that each participant experienced a range of music-story combinations that were widely spread across "emotion-space", they pose two difficulties for analysis. First, the analysis must assume that music extracts or stories within a category are effectively interchangeable; given that each category contained stimuli whose positions along either the movingness or the valence dimension may have been reasonably

[22]cf Winer (1971); cited in the *Statistica* Reference Manual.

disparate (look back at Figure 7.1 and Figure 7.2), this assumption is not necessarily valid; the lack of discrimination between disparate means leads to loss of analytical sensitivity. The second, related problem is that whilst an ANOVA can tell us, say, that T_TYPE had an effect of VL_MUS, it cannot shed any light on what *aspect* of the stories were responsible for the effect (movingness? valence? both?); this leads to lack of explanatory power.

In an attempt to circumvent these limitations, it was decided to re-analyse the data, this time ignoring the M_TYPE and T_TYPE clustering. Instead the specific extracts and stories presented to each participant in each trial were identified (using MUSIC and TEXT) and the results from Part One of the experiment replaced M_TYPE and T_TYPE as independent variables in a Multiple Linear Regression analysis that sought to test the power of the Part One variables MV_MUS, VL_MUS, MV_TXT and VL_TXT as predictors of the Part Two results.

Multiple Linear Regression

Five Multiple Linear Regressions were performed, one for each of the dependents MV_MUS, VL_MUS, MV_TXT, VL_TXT and SEP. As all took identical form, the data preparation and analysis procedure will be discussed globally here. The regression was not concerned with response profiles of individual participants, only the distribution of responses overall; therefore, data from all participants were concatenated to make just five columns of results, one for each dependent variable; each column contained 144 rows (12 trials x 12 participants = 144), each of which represented a single experiment trial answered by a single participant. Regression analysis requires that each datapoint is matched by an independent value with which it can be compared, so the next stage was to construct the independent variables for the regression. First, a single, independent MV_MUS and VL_MUS value was calculated for each music extract from the mean value scored by that extract in Part One of the experiment; the same was done with the stories to create a single MV_TXT and VL_TXT value. This effectively created for each extract and story a pair of values that corresponded to that extract or story's independently calculated position within "emotion-space". For each participant-trial (row), independent measures for MV_MUS, VL_MUS, MV_TXT and VL_TXT were assigned by using MUSIC and TEXT to look up which specific music extract and story was used in that trial, and assigning the relevant scores from the mean Part One results. Whilst this procedure may seem slightly convoluted, the effect was simple: it allowed analyses to be carried out in which the crude M_TYPE and T_TYPE independent variables were replaced with parametric equivalents that could both differentiate between trials based on the nature of the specific extracts and texts presented, and discriminate independently between dimensions of movingness and valence.

In the analyses that follow, the *dependent* variables will continue to be referred to as MV_MUS, VL_MUS, MV_TXT, VL_TXT and SEP; the *independent* variables (the mean Part One results for each extract and story) will be referred to as IMV_MUS, IVL_MUS, IMV_TXT and IVL_TXT.

Overall, there was no significant multiple regression for MV_MUS. Further examination of the effect of each independent, however, reveals that IMV_MUS was a significant predictor of scores ($\beta = 0.17$; $t_{(139)} = 2.05$; $p < 0.04$), and IVL_MUS showed a marginal negative regression at $\beta = -0.14$; $t_{(139)} = -1.66$; $p < 0.09$.[23] The effect of IMV_MUS merely implies that participants in Part Two of the experiment broadly agreed with those in Part One on the movingness of each music extract; the marginal negative regression with IVL_MUS suggests that participants found negative extracts more moving than the positive ones. Notably, IMV_TXT and IVL_TXT did not even come close to being predictors of MV_MUS scores. On this evidence, we cannot accept a hypothesis that reading a story affects the extent to which simultaneously presented music is found to be moving.

A multiple regression on dependent VL_MUS turned out to be extremely highly significant: Multiple $R^2 = 0.376$; $F_{(4,139)} = 20.93$; $p < 0.0000001$. This effect can be attributed to the predictive power of IVL_MUS ($\beta = 0.57$; $t_{(139)} = 8.45$; $p < 0.0000001$) and IMV_TXT ($\beta = -0.20$; $t_{(139)} = -2.68$; $p < 0.008$). The power of IVL_MUS is not interesting, as it simply points to agreement between Part One and Part Two participants. By contrast, the highly significant negative regression with IMV_TXT sheds light on the finding concerning T_TYPE in the ANOVA (see page 144): we can now see that the more moving a story was found in Part One of the experiment, the more negative the music seemed in Part Two. One explanation might be merely that the most moving stories happened to be those that were negatively valenced, and these would tend to make negative music seem more so, to "capture" ambiguous music, and to make people react negatively to the inappropriateness of music that would in another context evoke positive responses. But on the basis of this evidence, we can

[23] These values must, of course, be treated with extreme caution, because they are derived from analysis of a non-significant regression. They have been reported here, however, because they are nonetheless interesting.

accept a hypothesis that reading a story affects the valence of response to simultaneously presented music.

A highly significant multiple regression on MV_TXT (Multiple $R^2 = 0.17$; $F_{(4,139)} = 7.22$; $p < 0.000026$) can be explained mainly by the predictive power of IMV_TXT ($\beta = 0.42$; $t_{(139)} = 4.89$; $p < 0.000003$), indicating that for the third time, participants broadly agreed on ratings in Part One and Part Two of the experiment. However, IMV_MUS also contributed weakly to the prediction: $\beta = 0.13$; $t_{(139)} = 1.71$; $p < 0.089$. In other words, the findings tentatively support a hypothesis that the extent to which a story is found to be moving is affected by the extent to which a simultaneously presented music extract is found to be moving, a finding which film composers should find greatly relieving!

There was a highly significant multiple regression involving VL_TXT (Multiple $R^2 = 0.31$; $F_{(4,139)} = 15.59$; $p < 0.0000001$), all of whose predictive power can be attributed to IVL_TXT ($\beta = 0.53$; $t_{(139)} = 6.70$; $p < 0.000001$); none of the other independents were significant. These results, then, do not contradict a hypothesis that perceived valence of stories, by their linguistic nature conceptually specific, remain unaffected by simultaneously presented music.

Finally, a multiple regression analysing SEP was highly significant: Multiple $R^2 = 0.12$; $F_{(4,139)} = 4.86$; $p < 0.000004$. The main contributor was IVL_MUS ($\beta = 0.30$; $t_{(139)} = 3.66$; $p < 0.00036$); also, IMV_MUS was a marginal contributor ($\beta = -0.16$; $t_{(139)} = -1.80$; $p < 0.07$). This clarifies the related ANOVA results: negatively valenced, moving music was more likely to be integrated with stories than positive music. The implication of this finding is that people are more likely to integrate a music-story complex if the music is miserable! Perhaps a more plausible explanation is that moving music tended to integrate more with the stories that its less moving counterparts; the most highly moving music was negatively valenced.

Despite the strong evidence that ratings for movingness and valence of music and stories in Part One of the experiment could be good predictors for Part Two results, the Multiple R^2 values in all the analyses reveal a comparatively large Residual Variance that cannot be accounted for. Analysis of Residuals revealed that none fell outside $\sigma \pm 3$, indicating that there were no obvious outliers obscuring results. The unaccounted for variance must be seen, therefore, as an inevitable result of an experiment using large chunks of music, "real", complex stories and a relatively small number of participants. That trends emerge despite this is highly encouraging.

Secondary Variables

Data gathered from this part of the experiment would allow for the analysis of effects of several variables that have not yet been mentioned in this analysis, namely RB_MUS, RB_TXT, COMPL, REHEARD, AGE, HRSLIST, and HRSREAD. It was decided, however, not to investigate any of these variables further for two reasons: first, given the complexity of the analyses already undertaken, involving, as they do, transformations of data from Part One of the experiment as well as from Part Two, there appeared to be no obvious way of performing analyses in a way that would avoid the need to make arbitrary, statistically questionnable judgements concerning statistical approach; second, it was felt that analysis of these variables could not yield any findings that would shed further light on the primary purpose of this experiment, namely to investigate the effect of explicit external narratives on emotional response to music.

General Discussion

The results of this experiment are somewhat ambiguous, but they nonetheless indicate that listeners' reported emotional responses to music are affected by a simultaneously presented short story; to the extent that we can trust the validity of participants' introspective judgements, we can conclude that emotional response to music is affected by a simultaneously presented narrative. Further, the pattern of asymmetries found—response to a story affected the valence of response to music but not vice versa; response to music affected level of movingness in response to a story but not vice versa—is highly revealing and conforms well to the initial hypothesis set out at the start of this chapter: if the stories had simply acted as part of the context in which the music was heard and affected the listening process through mood induction (see Chapter Four) we should have expected valence and movingness ratings of the music to be affected equally; the fact that the stories had a significant effect only on the reported *valence* of the music suggests that propositional elements from the stories were being integrated into a "narrative" of the listening process. It should not surprise us to find, by contrast, (the unhypothesized results) that reported movingness of the music affected reported movingness of the stories whilst leaving the valence unaffected: the stories were self-sufficient narratives in their own right, so readers had no need to integrate elements of music into their narrative structures; that movingness levels were affected perhaps indicates that music provided a mood-moderating context within which the stories were read.[24]

The results of this experiment were obscured in part by a number of methodological problems. An immediate problem is that the experiment relied on participants' ability to introspect and report on their responses to the listening process as distinct to their responses to the reading process. Whilst introspective methods and seven-point rating scales are used regularly by music psychologists for many purposes, we should be wary of introspective judgements (see e.g. Nisbitt and Wilson, 1977), especially when applied to emotion. Although all participants claimed to be able to complete the ratings with no difficulty, a follow-up experiment using a different experimental approach that did not call for introspection would be highly desirable. It is quite possible that despite protestations to the contrary, participants were not, in fact, able accurately to report on their responses to the two domains separately, and that cross-contamination blurred results.[25] Another shortcoming of the use of rating scales in this context is their demand that the entirety of a dynamic listening experience be collapsed to a single, static number; even if participants had no difficulty performing this task, a vast amount of potential data is clearly lost by any objective criteria. A future investigation could perhaps take a similar experiment paradigm but substitute rating scales with a form of continuous measurement (see Schubert, 2000, and also Chapter Eight of this thesis).

Perhaps the biggest problem of all concerned the presentation of the stimuli: not only did participants differ significantly in terms of reading speed, but many found it difficult to listen and read simultaneously. This was evidenced both by informal comments and, to an extent, by the SEP ratings, which were somewhat higher than might have been predicted by the other data. Participants clearly were able to integrate response to music with response to the narrative—otherwise the pattern of results found in the other regressions could not

[24]In other words, the music was for the stories what a soundtrack is for a film; although these particular sound tracks were not necessarily suited at all well to the "films"!

[25]It could perhaps be argued that the potential difficulty of the task may have work to experimental advantage: participants attempted to subtract away their apparent response to the stories from their response to the music; that the stories still affected ratings of response to music indicates a powerful effect (of narrative integration?) whose effects were not open to introspection.

have emerged—but the task perhaps limited the extent to which they were doing so. Reading twelve short stories was hard work—it certainly required more effort than listening to the music—and the dataset may have been clearer if the task had been more passive (for example, if the stories been presented aurally). The original rationale for asking participants to read the stories themselves was three-fold: first, had stories been presented aurally, the reader's interpretation might have had a powerful effect on emotion judgements, not to mention the potential interaction between a dramatic voice and the music listening experience; second, the lack of consistent synchronization afforded by participants' differing reading speeds helped ensure that results were not coloured by freakish coincidence of any particular moment of music with a particular point in a story; third, forcing participants to indulge in a relatively active task, namely reading, increased the likelihood that they would remain focused throughout a long experiment session. A future experiment should perhaps substitute story reading with the presentation of short, silent films; this would eliminate differences in reading speeds at the expense of allowing potential coincidences to colour results. Listening to music whilst watching a film is also an activity with which most people are very familiar; it should be expected that perceived integration between domains would be stronger with film than they were with the stories, perhaps leading to more clear cut results all round.

A final problem with the experiment was simply that the task was too long; regardless of efforts to minimize the boredom factor, participants—some of whom who took almost two hours to complete the task—found the experiment sessions fatiguing and reported losing concentration.

Despite the difficulties, this experiment tentatively supports the hypothesis that we listen to music as narrative, or at least that the notion of narrative process provides a good explanation for some aspects of the listening process; it also provides further evidence of the oft-mooted observation that the process of listening and responding to music involves the integration of context. Perhaps more importantly in the context of this thesis, the experiment illustrates that to consider empirically the notion that responding to music can be attributed, in part, to narrative process is not only plausible, but could also provide a useful way of understanding emotional response to music in a multimedia—*i.e.* real—context. As has already been noted, virtually no work appears to have investigated the relationship between music and narrative in the evocation of emotional response to music despite the fact that such pairings are crucial to songs; where work has been carried out, it has mostly been in the context of music's effect on the emotion-evoking properties of film (see Chapter Five). There is one exception worthy of mention here, however, namely an investigation by Gfeller, Asmus, and Eckert (1991). In this study, participants' mood was assessed on a series of Affective Differential scales both before and after hearing a taped recitation of the poem *Barbed Wire* by R.H. Sauter; in various conditions, the recitation was heard alone or accompanied by either "commercial background" or "chromatic atonal" music. The investigation showed that "music can indeed alter affective response to verbal information"(p.140); sadly, despite its promise, the experimenters' decision to use only a single piece of text—a poem—presented auditarily, alongside only two musical extracts both of which were composed specifically for the experiment, prevents us from drawing any meaningful conclusions as to the nature of the relationship between the two domains.[26] If nothing else, it is hoped that the present experiment, in providing a second exception, will spark at least some interest in the notion that emotional response to music can best be understood in terms of a listener and his or her engagement with an environment bound together through some sort of integrative narrative process.

[26]The authors, clearly influenced by Berlyne (1971), did point to a relationship between complexity of stimulus (atonal vs. commercial music), mood and liking, but the paucity of stimulus material in conjunction with methodological shortcomings of the experiment preclude these being interpreted usefully within the present context.

Chapter 8

Anomalous Suspense and Factual Recall

Perhaps the most compelling conclusion of the experiment reported in the last chapter is that listeners appear readily to bind music and narrative into a single coherent experience, even when performing a task which is hardly optimised to maximise the likelihood of such binding; they also appear to be unaware of the extent to which they are doing so. An unsurprising corollary is that a listener's emotional experience in response to music is affected by a simultaneously presented narrative, although ambiguity in the results of the experiment—possibly caused by a number of methodological difficulties—preclude the drawing of firm conclusions regarding the exact role of narrative in the evocation of emotional response. One obvious way to proceed from here would be to refine the experiment along the lines suggested in the previous chapter: for example, stimuli could be presented aurally, or stories could be replaced with filmic narratives. The revised experiment, although similar in form to the original, would engage participants in a more dynamic multimedia experience, hopefully leading to a dataset with a considerably reduced noise floor and clearer findings. Given the time and length constraints of this thesis, however, and the wish to exemplify a variety of ways in which music as narrative might be researched empirically, it seems more appropriate to take a somewhat different approach. Rather than attempting produce a refined version of the same experiment, this chapter reports on a new design that seeks to corroborate the findings from a very different perspective. It seeks to demonstrate that the reason why musically-evoked narrative experience binds so readily with extra-musical narrative experience is that the two are really different instances of exactly the same mental phenomenon; further, it seeks to show that the coherent, dynamic nature of emotional response to music can be largely attributed to a listener's activity as a traveller to a Gerrigian narrative world. It will do so by trying to demonstrate an equivalence between response to music and response to textual narrative.

Possibly the biggest problem in designing an experiment intended to show equivalence of narrative experience across domains is the difficulty of isolating the phenomena under study: as was suggested at the end of the previous chapter, participants may well find it difficult to judge separately the effects of stimuli whose very co-existence encourages integration. In order to avoid this problem here, it was decided that this experiment would shirk simultaneous presentation of musical and textual narratives and, perhaps more importantly, would find a paradigm that did not require participants to introspect directly on narrative experience, or even on emotional response. Following the creed set out at the end of Part One of this thesis, inspiration for the paradigm was sought in the narrative psychology literature beyond the musical domain; it was found in a series of experiments carried out by Gerrig (1989).

Gerrig's experiments were intended to demonstrate that *anomalous suspense*, far from being a phenomenon whose manifestations are restricted to and explicable within the discourse of listeners' engagements with fiction, is in fact equally evident outside of the fictional domain; in addition, they were intended to shed light on the effect of prior knowledge on narrative experience. In the first experiment, which has already been described briefly in Chapter Five, participants were asked to read a series of very short stories (approximately 8 lines long) each of which took as its subject a well-known historical fact. Some of these stories merely gave a terse account of the events leading up to the fact in question; others were suspenseful, giving an account that in some way suggested a counter-factual outcome, casting doubt on the event in question. For example, one of the well-known facts was that George Washington was first President of the USA. A suspense-free story about this fact asserted Washington's popularity, his leadership ability and, in short, his eminent suitability for the post; a suspenseful one, by contrast, emphasised the man's age, his frailty, and his unwillingness to take up the Presidency. Both types of story ended with a "target sentence" that embodied either the factual outcome of the narrative (e.g. "George Washington was elected first president of the United States") or a counter-factual alternative generated by negation (e.g. "George Washington was not elected first president of the United States"); for each trial, participants had to state whether the target sentence was true or false. Gerrig hoped to find that despite prior knowledge of the facts, participants would feel anomalous suspense whilst reading the short stories and would become immersed in a narrative to the extent that even their real-world judgements would be impaired. Analysis of response times showed that this was indeed the case: participants were significantly slower at determining the validity of the target sentences when they when preceded by a suspenseful story.

Two follow-up experiments were designed to investigate the durability and robustness of the effect.[1] In the first of these, participants read each story as far as but excluding the target sentence. After reading all of the stories, they embarked on a five-minute distractor task, after which a series of statements was presented whose validity had to be judged; this series included—but was not limited to—the target sentences from the first part of the experiment.[2] Participants were led to believe that the reading of the stories and the verification of target statements were two utterly independent experiments: this deception was achieved by the invention of a new "task" that participants had to accomplish immediately after reading each story, namely, to think of a suitable title for it. The length of time taken to think of a title was ostensibly judged. Analysis of verification times showed that although the effect was somewhat weaker than in Experiment One, readers still took longer to decide whether a statement was true if it had been the subject of a suspenseful narrative.

The final experiment in the series eliminated the possibility that the results may have been caused not so much by the suspensefulness of a story, but by the extra information that participants had to consider when being presented with a story that implied a counterfactual outcome of an event with which they were familiar. This experiment was identical to Experiment Two in every respect, except that it introduced a new story category, "resolved suspense"; stories in this category were equivalent to the "suspenseful" ones in terms of both information content and in congruence to the real-world fact, but these stories resolved the

[1]These follow-up experiments were also intended to address some methodological difficulties that had emerged during the first, an attempt which the present author believes to be only partially successful, as several outstanding issues remain. These issues, however, are of no direct concern to us here, and will not be discussed further.

[2]Statements concerning facts that had not been the subject of any stories were added here to address a methodological problem of the first experiment, namely that all statements including the word "not" were false, whereas those without a "not" were all true. The extra true statements in this second experiment all included a "not", whereas the false ones did not.

suspense before the end so that congruence with reality was restored. As hypothesized by Gerrig, "resolved suspense" stories evoked validation response times that were not distinguishable from their "no suspense" counterparts; "suspenseful" stories continued to evoke slower responses. In other words, the degradation in response speed really could be attributed to anomalous suspense.

Although the details of Gerrig's experiment results and specifics of methodology used to gather them are of no particular relevance here, both the conclusions that he draws and his general methodological approach are of direct relevance. In concluding that "uncertainty [with regard to a well-know fact] can be induced by immersing readers in story episodes" (p.639) and demonstrating "how readily readers become immersed in stories"(p.646), Gerrig lends strong support to his theory (discussed in Chapter Five of this thesis) that anomalous suspense is the inevitable consequence of a narrative process that operates with an *expectation of uniqueness*. In other words, anomalous suspense, far from being a domain-specific phenomenon, is the hallmark of the mental processes involved in reading something—perhaps also hearing something—as narrative. Moreover, these experiments present a paradigm that can detect the symptoms of anomalous suspense—and, by extension, the presence of narrative process—without relying on introspective judgements. If it were possible to perform an equivalent experiment with music, demonstrating that the process of listening to music could, operating cross-modally, evoke anomalous suspense with respect to well-known facts in the same way that reading short stories appears to do and leading to the same pattern of response-time effects, this would constitute persuasive evidence that listening to music does indeed involve narrative experience, probably involving the exact same processes as those involved in reading a textual narrative. Rather than pointing merely to a crudely-diagnosed interaction between text and music, such an experiment could show that music and text are practically interchangeable in terms of their effect on narrative construction . This, in turn, would lend credence to one of the main propositions of this thesis, namely, that the dynamics and coherence of musically-evoked emotion are due, at least in part, to the fact that to listen to music is to be transported to a narrative world. The experiment discussed in the remainder of this chapter is an embryonic attempt to do exactly this.

Experiment

The experiments reviewed above demonstrate that uncertainty created by exposure to a suspenseful short story about a well-known historical fact affects the speed with which a reader is able to assess the validity of that fact when subsequently presented. Our experiment here is designed to test the hypothesis that uncertainty created by exposure to a suspenseful musical extract about a well-known historical fact will affect the speed with which a reader is able to assess the validity of that fact when subsequently presented.

Macroscopically, an experiment to do this could take on a form very similar to the second experiment reviewed above: participants could be asked to listen to a series of suspenseful musical extracts "about" well-known historical facts, after which they could be asked to assess the validity of a series of factual or counter-factual statements. The more suspenseful the music, the longer it should take participants to judge the validity of the fact to which it pertained. Immediately, however, two rather serious problems present themselves: first, we would need some way of encouraging participants to relate the music to well-known historical facts, a problem that Gerrig did not face with his clearly relevant stories; second, we would need to find some way of classifying the music as suspenseful or otherwise, a task that would probably involve complex criteria and subjective judgements, leading to categorisations far less robust than those deriving from Gerrig's clear and straightforward textual classifications. Fortunately, both of these issues can be resolved.

In order to forge a relationship between a well-known historical fact and an arbitrarily assigned musical excerpt, we can make use both of the findings of the experiment reported in Chapter Seven—listeners readily bind music with simultaneously presented text—and also evidence cited in Chapters Four and Five, which suggested that central to any listening experience is a web of associations, musical, situational, and historical, that form the context in which music is heard. Listeners will happily relate a musical extract with a well-known historical fact and integrate that relationship into the narrative of a listening experience so long as they are given an excuse to do so; a trick as simple as asking them to imagine that a musical extract is *intended* to reflect the events leading up to a well-known fact should be enough to encourage most listeners to entertain the notion that the music really is "about" those events. If this hypothesis is correct, any suspensefulness detected in the sound or structure of the music—listeners' responses to sound, utterance and context—will cohere into a narrative that should be able to replace Gerrig's fictional stories. What criteria should we employ for choosing musical extracts for the experiment? According to these arguments, it should not matter very much; the only important criterion is that each extract could plausibly be considered to be "about" the relevant historical events; participants might have trouble integrating the Battle of Banockburn with a xylophone orchestra playing *Walzing Matilda*; the incongruence would most likely lead to a context whose salient attribute is not so much battle but surrealism!

Whatever the nature of the musical extracts chosen as stimuli, a music-based adaptation of Gerrig's experiment requires that these extracts be categorised as suspenseful, suspenseful with resolution, or non-suspenseful. Given that *ad hoc* categorisations made by the experimenter might be considered somewhat idiosyncratic, it would perhaps be more appropriate to take the same approach to categorisation as that taken in the experiment reported in Chapter Seven of this thesis, namely to delegate the task to a group of independent listeners. Even with this approach, however, a problem remains: the categorisation of musical extracts into three distinct groups is not an easy task; it is notoriously difficult unambiguously to categorise musical phenomena on any dimension, let alone one with as much potential for idiosyncrasy as suspense evocation. Perhaps a more appropriate approach would be to ask the judges to attempt the comparatively easy task of rating the suspense level along a parametric scale; although the statistical design of the experiment would change—we would be looking for a regression of response times against suspense level rather than different means across categories—the fundamental meaning would be identical. Even this solution leaves a little to be desired, however, because we are not interested in how much suspense a typical listener might report detecting in a given piece of music, something which is likely to a function of recognition of convention rather than anything else; instead, we need to know how each specific piece of music *when associated with the events leading up to a well-known fact* evokes suspense in a particular listener leading to uncertainty with respect to the outcome of those events. The only way to achieve this is to ask the listeners who are taking part in the experiment to rate levels of suspense themselves as they are subjected to each trial. If a listener, immediately after each trial, rates the suspensefulness of the listening experience, such ratings could be used safely as predictors of response times in a regression analysis.

One of the most impressive features of Gerrig's series of experiments is that it did not ask participants to make any introspective judgements; reading the stories and assessing the validity of the statements was all that was needed. By contrast, the task suggested for the present experiment requires participants to make subjective judgements that are retrospectively introspective; they would be asked how much suspense they felt (introspection) whilst they were listening (retrospectively). To be sure, it is highly unlikely that participants would relate the first part of the experiment (in which they would be asked to rate suspense) with the second more objective part (in which they would assess the validity of a series of statements) so the experiment paradigm remains fundamentally robust, but it would be far better if reliance on introspection were minimized. After all, as has often been noted, participants are notoriously bad at rating their responses to musical stimuli, especially where judgements related to emotion are concerned. Part of the problem, of course, is that a listening experience is dynamic and extended over time; it is hardly suprising that listeners have a problem reducing such an experience to a single number on a seven-point scale.

In recent years, some music psychologists studying emotion have attempted to address this problem by using on-line data gathering techniques (e.g. Nielsen, 1983; Madsen and Fredrickson, 1993), which replace seven-point rating scales with tools that allow subjects to judge the phenomenon under study in real time as the music plays. These tools involve participants moving a computer mouse, a joystick or some similar device through a one- or two-dimensional space, where each dimension represents one of the parameters under study. For example, had the experiment reported in Chapter Seven used an on-line tool, it might have asked participants to judge the texts and the music by moving a joystick through two-dimensional space, where one dimension was arousal and the other valence. Although many studies using such techniques suffer from the difficulties of analysing continuous data (see Schubert, 1998, 2000), they have proven very successful in capturing actual listening experiences not only because they ask listeners to respond in real time thereby obviating the need for retrospective judgement, but also because partipants must respond continuously as they listen, leading to drastically diminished scope for analytical introspection. In addition, participants appear to find these on-line response tasks rather natural and easy to perform compared with their rating-scale counterparts. This final observation is perhaps unsurprising given that the dynamic nature of the task matches well the dynamic nature of a listening experience.

An on-line response paradigm seems ideally suited to the needs of the present experiment: rather than asking participants to retrospectively introspect on the suspensefulness of the listening experience, they can be asked to track it in real time on a unidimensional scale. These continuous data can then be transformed into values appropriate for use as predictors in a regression analysis. Substitution of a computer mouse or joystick with a specially designed pressure sensor could make the task surprisingly natural: participants merely need to squeeze the sensor with a force that reflects the suspense they are feeling at any given moment; anyone who has ever watched a moderately suspenseful film at the cinema and has literally clung to the side of a chair will appreciate quite how natural a task squeezing a pressure sensor in response to suspense could be!

A combination of the on-line response paradigm with Gerrig's second experiment gives us a means to test the hypothesis that uncertainty created by exposure to a suspenseful musical extract about a well-known historical fact will affect the speed with which a reader is able to assess the validity of that fact when subsequently presented. This, in turn, will test the hypothesis that responding to suspense in music involves narrative processes that are at the very least homologous—if not identical—to those involved in responding to suspense in textual narrative.

Method

Participants were asked to listen to a series of musical extracts whilst entertaining the proposition that each extract was "about" the events leading up to a real historical event whose outcome is well known; during the listening process, they were also asked to squeeze a pressure sensor to reflect the level of suspense they were feeling. After each extract, they rated how well the music fitted the historical event.[3] Having heard all of the extracts, participants were presented with a series of factual or counter-factual statements, one about the outcome of each of the well-known historical events presented in the first part of the experiment. The task was to assess the validity of each statement by clicking a "true" or "false" button at a computer console; the time taken for each assessment was recorded.

[3]Although this measure was not of direct interest to the experimenter, it was felt that it might help in the interpretation of other findings. The measure also served a procedural purpose: its presence increased the gap between each trial, and lent more plausibility to the claim that this part of the experiment was unrelated to the final part (see Procedure below).

Stimuli: propositions and statements

20 historical events were taken as subject matter for the propositions and statements. Events were chosen according to the following three criteria:

- Circumstances leading up to each event should involve an element of uncertainty; in every case, there should have been at least a plausible possibility that the outcome may have turned out differently. This criterion was important for ensuring that a musical narrative "about" these circumstances had the potential to be suspenseful.
- Each event should involve an element of human interest; the outcome should have ramifications for a historical individual or society. Again, this criterion increased the potential for a suspenseful narrative.
- There should be a strong likelihood that each event would be familiar to all experiment participants, at least to the extent that every participant would be able to distinguish a factual from counter-factual outcome. Given that the experiment was concerned with detecting the effect of musical narratives on ability to recall real-world facts, this requirement was clearly essential!

Chosen subjects were all either well-known events in British history or Twentieth Century world history, concerned a well-known historical figure, or were recent happenings which have received significant news coverage within the last decade.

For each subject, a *proposition*, a *factual statement* and a *counter-factual statement* was created. The propositions would be presented to participants as the subject matter to consider during the playing of the musical extracts; it was important that they should facilitate suspenseful narrative formation, unambiguously relate to the events in question but not prejudice participant's prior understanding of the facts. To this end, they were formulated according to the following criteria:

- Experiment participants should immediately be able to identify the subject matter from the proposition alone.
- Each proposition should encapsulate antecedents of the event in question but not the outcome; they should evoke both the element of human interest inherent in each subject and the uncertainty.

- Propositions should be as neutral as possible with respect to the actual outcomes of events; they should avoid suggesting preferentially either the factual or counterfactual outcome of the events in question.
- The phrasing of each proposition should admit a single ground for uncertainty.

All propositions were written in the present tense; uncertainty and human interest were implied by the regular use of formulations such as "tries to", "attempts to", "decides whether to", etc. Where, in the judgement of the experimenter, propositions required more human interest or uncertainty, adverbs such as "desperately" were inserted in suitable locations.

Factual statements consisted of single sentences that encapsulated the actual outcome of the event alluded to in a proposition. Each factual statement was formulated to maximise the probability that experiment participants would immediately recognise the subject to which it referred, without having to refer back to the associated proposition. In contrast to the propositions, factual statements employed language that was as objective and unimpassioned as possible, simply stating an outcome as tersely as they could. Inasmuch as it did not conflict with these goals, factual statements used language that was as similar as possible to the propositions, to maximise the likelihood that participants would make a (perhaps subconscious) link back to the musical narrative when assessing the validity of the statement.

Counterfactual statements were formulated according to similar criteria to their factual counterparts except, of course, that they encapsulated a counterfactual alternative to the actual outcome of each event. An additional criterion for their formulation was that within the context of the experiment itself, each counterfactual outcome should be worded so as to seem equally as plausible as the factual version; in other words, regardless of the subject matter and the likelihood of counterfactual—or for that matter factual—outcomes actually occurring, the sentences should be structured in an equivalent way. In practice, this meant that in contrast to Gerrig's experiments, counterfactual statements avoided the use of negation, and instead used alternative formulations. To clarify, an example of a factual outcome (although in this case fictitious) might be "humpty-dumpty fell off the wall"; its counterfactual counterpart might be "humpty-dumpty remained seated on the wall" rather than "humpty-dumpty did not fall off the wall" because the latter statement implies its factual alternative.

Propositions, factual statements and counterfactual statements for each of the 20 subjects are listed below (propositions are marked with the letter "P", factual statements with "F" and counterfactual statements with "C"):

P. *The East German regime tries desperately to prevent the destruction of the Berlin Wall*

F. *With a crumbling East German regime, the Berlin Wall was finally knocked down*

C. *Despite a crumbling East German regime, the city of Berlin remained split by The Wall*

P. *Henry VIII decides whether to order the execution of Anne Boleyn*

F. *Henry VIII ordered the Execution of Anne Boleyn*

C. *Henry VIII ordered that Anne Boleyn should be pardoned and released*

P. *Neil Armstrong attempts to become the first American to walk on the moon*

F. *Neil Armstrong succeeded in walking on the moon*

C. *Neil Armstrong never managed to walk on the moon*

P. *After four consecutive terms of Conservative rule, the Labour Party tries to oust them from office*

F. *The Labour Party ousted the Conservatives from office in the 1997 elections*

C. *The Conservative Party managed to hold on to power in the 1997 elections*

P. *After hitting an iceberg, the Titanic liner tries desperately to reach dry land and safety*

F. *After hitting an iceberg, the Titanic liner sank with great loss of life*

C. *After hitting an iceberg, the Titanic liner managed to struggle to dry land, saving many lives*

P. *The War Cabinet, knowing that Coventry will be bombed, decides whether to evacuate the city and reveal to the Germans that they have cracked the Enigma code*

F. *Despite knowledge gained from cracking the enigma code, the War Cabinet refrained from warning Coventry of an impending German bomb attack during World War II, resulting in great loss of life*

C. *The War Cabinet used their knowledge gained from cracking the Enigma code to warn the people of Coventry of an impending German bombing raid during World War II*

P. *At the culmination of the Cuban Missile crisis, Krushchev decides whether to withdraw from Cuba or risk nuclear war*

F. *At the culmination of the Cuban missile crisis, Krushchev withdrew from Cuba averting nuclear war*

C. *At the culmination of the Cuban missile crisis, Krushchev stood firm, leading to nuclear war*

P. *President DeKlerk's government decides whether Nelson Mandela should be freed after more than two decades in captivity*

F. *President DeKlerk's government decided to free Nelson Mandela after more than two decades in captivity*

C. *President DeKlerk's government decided to keep Nelson Mandela imprisoned even after more than two decades in captivity*

P. *President Kennedy fights for his life after being shot by Lee Harvey Oswald*

F. *President Kennedy died after being shot by Lee Harvey Oswald*

C. *Kennedy survived after being shot by Lee Harvey Oswald*

P. *US Congress votes on whether to pass the Civil Rights Act, containing much needed anti-discrimination laws*

F. *US Congress voted to pass the Civil Rights Act, containing much needed anti-discrimination laws*

C. *US Congress voted to reject the Civil Rights Act, containing much needed anti-discrimination laws*

P. *A consortium of British and French aerospace companies try to design Concorde, a passenger jet that can fly faster than sound*

F. *A consortium of British and French aerospace companies made Concorde fly faster than sound*

C. *A consortium of British and French aerospace companies failed to make Concorde fly faster than sound*

P. *German police attempt to free eleven Israeli athletes who are being held hostage by armed terrorists at Munich airport*

F. *Eleven Israeli athletes being held hostage by armed terrorists in Munich all died after a failed attempt to save them*

C. *Eleven Israeli athletes being held hostage by armed terrorists in Munich were all res-*

cued and brought to safety

P. *President Allende of Chile attempts to survive a military coup*

F. *President Allende of Chile was overthrown in a military coup*

C. *President Allende of Chile survived an attempted military coup*

P. *After years of unrest and violence, the South African government finally decide whether to abolish apartheid*

F. The *South African government eventually abolished Apartheid*

C. The *South African government never abolished Apartheid*

P. *President Nixon decides whether to resign or face impeachment*

F. *President Nixon resigned to avoid impeachment*

C. *Having refused to resign, President Nixon was impeached*

P. The *Argentinian government decides whether to risk was with Britain and invade The Falkland Islands*

F. *Argentina invaded the Falkland Islands*

C. *Argentina dropped their plans to invade the Falkland Islands*

P. *Nuclear scientists desperately try to control the reactor at Chernobyl to prevent a disaster*

F. *Nuclear scientists failed to control the Chernobyl reactor, leading to disaster*

C. *Nuclear scientists managed to control the Chernobyl reactor, preventing disaster*

P. *The "Challenger" space shuttle attempts take-off, the most dangerous part of its mission*

F. *The "Challenger" space shuttle exploded soon after take-off, killing all astronauts on board*

C. *The "Challenger" space shuttle completed all its missions safely*

P. *A jury decides whether to convict L.A.P.D. Officers involved in the vicious beating of Rodney King*

F. *The L.A.P.D. Police Officers involved in the Rodney King beating were acquitted by the courts*

C. *The L.A.P.D. Police Officers involved in the Rodney King beating were convicted by the courts*

P. *Captain Scott attempts to return to the base camp safely in freezing temperatures after reaching the South Pole*

F. *Captain Scott died before reaching his base camp on his return from the South Pole*

C. *Captain Scott successfully reached the South Pole and returned home*

All the propositions were typed, in arbitrary order, into a text file in a format that could be read by the computer program which presented trials. The statements were split into two separate files, according to the following procedure: first, a random selection of 10 statements was put into file A and their counterfactual counterparts put into file B; next, the remaining 10 factual statements were appended to file B and the remaining 10 counterfactual statements to file A. The result was that each of files A and B contained exactly 10 factual statements and 10 counterfactual statements, and each file contained exactly 1 statement for each of the 20 propositions.[4] During experiment sessions, both propositions and statements were displayed on a computer screen in a large, easily legible font.

Stimuli: music

Twenty musical extracts were taken from real recordings (i.e. not MIDI) of film soundtracks. Such music is ideally suited to this type of experiment, because it tends to have a fragmentary structure that does not rely on large-scale architecture for coherence, and is meant to be heard as part of a multimedia scene. Extracts were chosen according to the following criteria:

- Each should have the potential to evoke some degree of suspense when paired with the statements.
- Each should be a plausible soundtrack to any of the statements. In practice, this criterion meant that they had to be reasonably ambiguous with respect to emotions represented (a feature that is not hard to find in the film music repertoire).
- Between them, the extracts should be varied enough to elicit a range of suspense reactions from participants but should be similar to each other in as many other respects

[4] Owing to a stimulus-preparation error that lurked unnoticed until the results of the experiment were analysed, file A really contained 11 statements; both the factual and counter-factual statement concerning Captain Scott were included. The effects of this oversight were corrected for prior to statistical analysis (see report of results below).

as possible, so that the "environment" created in an experiment session remained reasonably consistent throughout. It was hoped that this stipulation would maximise the likelihood that participants would be able to find and hold on to a consistent strategy for using the pressure sensor.

- The source of the extracts should not be recognisable; in other words, none should contain a theme tune or distinguishing mark that would allow participants to relate a trial in which it is heard to the film from which it originally came.

All the extracts chosen were from scores by Georges Auric (music from *Caesar and Cleopatra*, *Dead of Night*, *Father Brown*, and *It always rains on Sunday*), Patrick Doyle (music from *Henry V*) and John Williams (music from *Black Sunday* and *The Empire Strikes Back*). Details of the discs and tracks from which they were taken can be found in Appendix B.

All extracts were approximately 45 seconds long (although one was approximately 33 seconds long, and another lasted for approximately 1 minute). Having chosen appropriate recordings and tracks, the specific portion of music used for each extract and its exact length were selected according to the contingencies of the music itself: each extract started at a musically plausible starting moment and was cut off as near to 45 seconds later as possible whilst still stopping at a musically appropriate cut-off point.

The stimuli were prepared by loading the relevant CD tracks into a computer-based digital editor, choosing exact start and end points for each extract, and then applying fast fades and equalisation to make each extract sounded like a coherent whole, starting and ending smoothly. All 20 edited extracts were then converted to MP3 format (using the *xingtech* variadic bit-rate encoder) and loaded on to the computer on which the experiment would be run.

Stimuli: presentation in trials

Stimuli for each trial in the first part of the experiment consisted of presentation of one of the statements and one of the musical extracts. Participants were presented with each statement only once and heard each musical extract only once; the pairing of statements with extracts and the order in which they were presented was randomly generated for each participant by the experiment software (see below).

For the second part of the experiment, each of the propositions in *proposition file A* or *proposition file B* was presented, again in a random order determined for each participant by the experiment software. The proposition files were assigned to participants on an alternating basis, so that odd numbered participants were presented with the contents of *file A* and even numbered participants with *file B*. Given that the order in which participants attended the experiment session was merely a function of when each individual happened to have a spare half an hour, the assignment was effectively random, but ensured that the same number of participants were presented with the contents of each file.

Apparatus

Experiment sessions were controlled by a PC-based Linux workstation running software written specifically for the experiment, which was responsible for presenting stimuli and collecting responses.[5] Sound hardware consisted of a *Maestro 3i* card connected to Sony *MDR-V4* headphones; visual stimuli were presented on a 15″ colour LCD screen. In the final part of the experiment, responses were given by participants using a computer mouse to click on a "true" or "false" button; the software recorded both the response itself and the elapsed time between the statement being displayed and the button being clicked; tests during development revealed that measurements were accurate to within a few milliseconds.

Suspense ratings were collected using a pressure sensor designed and built specifically for the experiment.[6] The sensor was a small device intended to be held in the palm of a hand and squeezed with a thumb; it interfaced with the computer via the parallel port. Every pressure sample was measured as a value between 0 and 127, where 0 is the lowest measurable pressure and 127 the highest; the sensor was calibrated so that 0 corresponded to no pressure being applied, and 127 corresponded to a thumb grip that, whilst very firm, would have been easily within the power of all participants to achieve. Between the two extremes, the scale was approximately linear. Visual feedback in the form of a computer dial informed partipants how hard they were squeezing at any given time. During each trial, the pressure sensor's value was sampled 10 times per second, a sampling rate that was well within the capabilities of the hardware; sampling rate, data collection and the visual

[5] The full source code of this C++ program, along with a brief explanation of its operation is presented in Appendix A. Unfortunately, the need to interface with specialised hardware and record accurate timings meant that *The MuPsych Lab* could not be used for this experiment.

[6] A full description of its design, including a circuit diagram, can be found in Appendix A.

feedback were handled by the computer software.

Participants

17 people participated in the experiment. All were students or staff at the University of Cambridge. Three participants were music students (2 undergraduate, 1 postgraduate); none were psychologists. Approximately half of the participants had previously taken part in the experiment reported in Chapter Seven, but this was carried out almost one year prior to the present experiment so no cross-experiment contamination was to be expected. Given the lack of any interesting effects found in Chapter Seven relating to participant age or time spent listening to music, these data were not collected here. 8 participants were presented with set of facts A; 9 were presented with the complementary set B.

Procedure

On entering the experiment room, partipants were asked to sit in front of the computer and read the following on-screen instructions:

> *The first part of this experiment will consist of 20 trials. To complete these trials, you will need to use the* computer mouse, *the* rating sheet *(which may be found to your left) and the* pressure sensor *(which may be found to your right). Please locate these items now; the experimenter will show you how to use the pressure sensor.*

At this point, the experimenter demonstrated use of the pressure sensor and ensured that participants had located the mouse and rating sheet.

> *Each trial will take the following form:*
>
> *You will be presented with a sentence that describes a scenario; this scenario will set out the antecedents of a well-known historical event. (You may well recognise the scenario in question, but do not worry if this is not the case). Read the sentence and spend a few moments contemplating the scenario, its tensions, and ramifications of potential outcomes.*
>
> *When you are ready, click on the button marked "click here to start music", and a musical excerpt will play for approximately 40 seconds to 1 minute. Imagine that this music, just like a film score, has been written specifically to reflect the tensions and uncertainties inherent in the scenario, that its unfolding is meant to reflect the unfolding of the scenario.*

> *As the music plays, use the pressure sensor to "track" the level of suspense you feel with respect to this unfolding "scene". A dial at the bottom right-hand side of the screen will give you a visual indication of the pressure level you are applying to the sensor.*
>
> *Once the music has stopped, judge of on a scale of 1 - 7 (where 1 = not at all well; 7 = extremely well) how well you think the music reflected the tensions of the scenario; your judgement should be recorded on the rating sheet. Having made your judgement, click on the button at the bottom of the screen to move on to the next trial.*
>
> *When you are ready to start the experiment, put on the headphones and click the button at the bottom of the screen; the first scenario will then be displayed. The first two trials are for practice only and your answers will not be recorded. Use these two trials to familiarise yourself with the task, and do not hesitate to ask the experimenter if you are unsure about anything.*

Having read the instructions, the experimenter watched each participant complete the first two trials, and asked for confirmation after each of them that the participant fully understood the task. During the first trial, just before clicking to start the music, participants were encouraged to practice squeezing the pressure sensor, and to explore the full range of the scale and the way in which it reacted. Each participant was informed verbally: "try getting the dial to move all the way round the scale, so you can see what the full range feels like. You do not necessarily have to use the full range—use a range that feels comfortable—but it is as well to know what the full range actually is". Participants were left to complete the remaining 18 trials alone.

After completing the trials, participants were asked to remove their headphones and were given a break of approximately two minutes, during which time they were asked whether they had enjoyed the task. Many participants asked about the origins of the musical extracts; they were told merely that "they came from various film soundtracks". The experimenter then asked participants whether they would mind starting the second, separate, part of the experiment which would only take a few moments. The following instructions were displayed:

> *In this part of the experiment, you will be presented with a series of statements pertaining to well-known historical events. Your task is to judge whether each statement is true*

(factual) or false (counter-factual); you give your answer by clicking on the relevant button (markedly true *and* false *respectively), using the computer mouse. Statements will be presented one after the other in immediate succession; as soon as you click on* true *or* false, *the next one will appear. Judge each statement as quickly as you can, but without compromising the certainty of your judgements.*

Most participants spent between one and two minutes completing this part of the experiment, after which they were debriefed, and were encouraged to ask any questions about the tasks they had completed.

The entire experiment session lasted for approximately 25 minutes.

Design

The complexity of the data—a combination of rating scale, response times and continuous pressure data—preclude the possibility of a straightforward traditional design. The main analysis took the form of a Multiple Linear Regression whose aim was to predict response time with downsampled pressure data. Details are presented in the results section below.

Results and Discussion

The null hypothesis to be tested was that suspense in a music-proposition pairing (as encoded in a pressure trace) would not affect the speed at which participants were able to assess the validity of a fact related to a given proposition. Therefore, the main analysis consisted of two parts: first, data were transformed into a form in which pressure traces and response times could be directly compared; second, comparisons were performed using a Multiple Linear Regression procedure.

Prior to any data transformation or analysis, pressure traces from the first part of the experiment were coupled with the response times from the associated fact in the second part. Then, traces from the first two music-proposition pairings presented to each participant (i.e. the practice trials) were removed from the dataset, along with the response time data from the associated facts. Although a result of this procedure was that not all propositions were represented by the same number of datapoints, all trials had been randomly ordered for each participant, so the potential for serious imbalance was minimized.

Preliminary examination of the dataset revealed an error in preparing the stimuli that resulted in participants presented with fact set A being asked to assess the validity of two facts about Captain Scott (both the factual and counter-factual statements were presented). It was decided, therefore, to remove response time data from the *second* question presented to each. Although this led to an uneven number of responses to factual and counterfactual statements about Captain Scott's fate, given the size of the dataset overall, this was felt to be unimportant.

These procedures left 18 trials per participant for analysis.

Data Preparation

As the hypothesis concerned the effect of suspense on the recall of *well-known* facts, the first stage in the analysis was to ensure that the facts presented really were well known. Sadly, they were not quite as well-known as had been predicted! Although three participants correctly identified all 18 of the facts presented to them, four correctly identified only 13 or fewer. The mean score was 15.65 correct validations with a standard deviation of 2.00.[7] All response times for incorrectly validated statements were removed from the dataset, along with their associated pressure traces. This procedure, in conjunction with the removal of practice trials and compensation for the Captain Scott problem, left 266 pairs of response times and associated traces for further analysis.

Response times were distributed with a mean of 3968ms and a standard deviation of 2728ms. The fastest response took a mere 57ms and the slowest 24001ms. Given the impossibility of someone reading a fact and responding within 57 milliseconds, we can assume that this outlier was a result of the participant accidentally double-clicking the mousebutton, thereby both popping up a fact and responding to it in one motion. As for the 24 second response, it seems safe to assume that the participant either had no clue as to the validity of the statement, or else experienced a serious concentration lapse. These outliers were not removed from the analysis at this stage.[8]

Informal observation of the experiment sessions suggested that participants differed

[7] Standard deviation is not a particularly informative measure here, as the spread of scores did not approximate a Normal distribution. Further, it should be noted that these statistics constitute an optimistic view of how well the facts were known, as we can expect that participants who guessed responses had a 50% chance of guessing correctly.

[8] In fact, they were not removed at all, as they were found to have negligible impact on later statistical analyses.

wildly in the average time it took them to assess the validity of facts. Mean response times per participant were indeed highly varied: mean mean response time was 4048ms with a standard deviation of 1736ms; the fastest mean was 2227ms and the slowest 8019ms. In order to avoid participant response-time profiles becoming a major factor in analyses, it was decided to normalize response times *per participant* prior to any further analysis. A Pearson Product-Moment one-tailed test revealed no significant correlation between participants' mean response times and the number of correctly judged statements.

Preparation of the pressure trace data was, by necessity, a slightly more idiosyncratic process. Each trace consisted of between approximately 300 and 600 numbers, each representing the pressure being applied to the sensor in a 100ms sample period; the goal was to transform these data into a form that would (a) allow direct comparison across traces, (b) remain a good reflection of recorded suspense level and (c) allow comparison to be made with response time data using standard statistical techniques. The first requirement dictated that whatever transformation was used, the resulting dataset would have an equal number of values for each case, and the second that the resulting dataset would preserve pressure levels as recorded by the original trace. The third requirement suggested an approach that reduced the number of data points as much as possible without violating the second. Given that informal observation of the data—and comments from some participants—suggested that high-frequency changes could be attributed mostly to unintended variations in pressure level, it was decided to transform the traces with a *downsampling* procedure.

The purpose of the procedure was to transform a trace into a small number of data points, each of which accurately reflected the average pressure over a range of samples; this small number of points could then be tested with a Multiple Linear Regression analysis to see whether average level of suspense at various points during a trial was a good predictor of response times. It was felt that downsampling in this way constituted a satisfactory compromise between preserving the original form of the traces and yielding a manageable data set from which meaningful data could be extracted. As the rationale for using a pressure sensor in the first place was to avoid the retrospective introspection of a rating scale, the fact that accurate temporal information would be lost during a downsampling operation was considered unimportant.

Traces were downsampled to 5 points, each of which represented the average pressure—

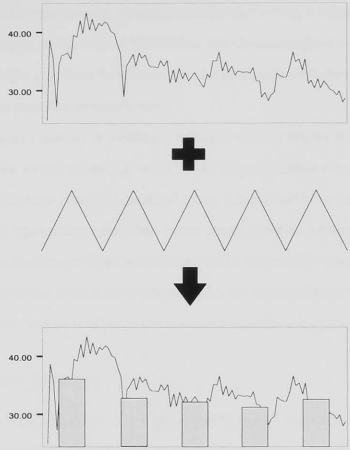

Figure 8.1: Downsampling with a triangle function. The top graph is an original pressure trace (chosen at random from the dataset). Its downsampled representation can be seen in the bottom bar graph. Note that the downsampled values are a good reflection of the overall profile of the original graph despite the lack of resolution and high frequency components.

and by extension average suspense—over approximately a 10 second period.[9] Values were calculated by dividing a trace into 5 equal sections and taking a weighted mean value for each section. The relative weighting given to each point was determined by a *triangle function*: an equilateral triangle was overlaid on the trace and the value at each point multiplied by the relative height of the triangle at that point (see figure 8.1). The effect of the triangle function was to attach more weight to data points in the middle of each sample than to those at the edges; in other words, it sharpened the distinction between the points, ensuring that low-frequency variation in the original traces would remain differentiated in the downsampled values.[10]

After downsampling, each trace was represented by 5 datapoints; differences between

[9]Visual observation of a random sample of pressure traces in conjunction with pragmatic concern for a "good" number of data points with a midpoint suggested that 5 samples was a sensible number to choose.

[10]The author would like to thank Alistair Turnbull (Computer Laboratories, Cambridge University) for suggesting the use of a triangle function here, and providing an algorithm for calculating the values.

the values of the points represented change in suspense during a listening experience. The final stage in data preparation was to normalize the downsampled values per participant for consistency with the response time data and to minimize variation due to participants' differential use of the pressure sensor's range.

The above processes resulted in a dataset containing 266 rows, each of which represented 5 samples of suspense rating recorded whilst a participant listened to a musical excerpt in the context of a proposition about a historical event whose outcome was known to the participant, and a value representing the time taken for that participant subsequently to judge the validity of a statement regarding the outcome of the historical event in question. All data processing described in this section was performed with a computer program written for the purpose; this program, with a brief explanation of its function, is presented in Appendix A.

Multiple Regression Analyses

In order to test whether suspense was a good predictor of response times, the data were subjected to a Multiple Linear Regression analysis. A stepwise procedure was chosen, using a cutoff of $p < 0.05$ for entering potential predictors and a cutoff of $p < 0.10$ for removing them. Independent variables for the analysis were the downsampled values representing average suspense level at various points along the pressure trace (SAMPLE1, SAMPLE2, SAMPLE3, SAMPLE4, SAMPLE5); the dependent variable was response time (RESPTIME).[11]

Analysis revealed no significant regression whatsoever; pressure level as represented by the five downsampled values was completely unable to predict response time. Attempts to change the number of downsampled values (3, 4 and 6 were tried) also revealed no significant results. These data, then, are unable to confirm that suspense evoked by a listening experience affects the speed at which people are able to judge the validity of a fact concerning the outcome of a well-known historical event; the null hypothesis cannot be rejected.

There are several possible explanations for this result. First, the premise on which the hypothesis was based may not be valid: perhaps the experience of listening to music does not involve processes homologous to the narrative processing evident in comprehension of texts and films. Second, the results of Gerrig's experiment (upon which the present experiment was based) may not be robust; if this is the case, we have no reason to assume that

[11] All statistical analyses were performed with SPSS v.10

narrative processing should affect the speed at which someone is able to judge the validity of a well-known fact. Third, the pressure traces—and therefore the downsampled predictor values—may not constitute an accurate measure of suspense; participants may have been unable to use the sensor in the requested way.

In the light of evidence presented earlier in this thesis and elsewhere, the first suggestion seems highly unlikely; at least, the notion that music is heard as narrative should certainly not be rejected on the basis of this experiment alone! The second is somewhat more likely, as Gerrig himself noted that results of his second experiment (in which presentation of stories was not followed immediately by judgement of response times) were rather weaker than results of the first (in which each story was immediately followed by the judgement task). A replication of Gerrig's experiments, or possibly a controlled combination of them and the present one, could help resolve this issue. Most likely of all, perhaps, is the third possibility: asking participants to rate suspense continuously by applying pressure to a sensor certainly avoided problems associated with retrospective judgements, but it still required introspection and real-time introspection at that! It is perfectly possible that participants were simply unable to listen to the musical excerpts whilst entertaining the linked proposition *and* provide a continuous pressure trace reflecting level of suspense. It was suggested earlier in this chapter that squeezing a pressure sensor in response to a suspenseful moment in a narrative is a very natural thing to do; in retrospect, it seems plausible that tracking suspense consciously and continuously over a period of 40 seconds or more with a pressure sensor is not quite such a natural thing to do.

Even if level of suspense is not reflected directly in mean pressure level, however, it may well be encoded in some other aspect of the pressure trace less overtly under participants' conscious control. One obvious possibility is that amount of *change* in pressure level at any given sample period may be a better reflection of perceived suspense. We could perhaps hypothesize that wide variation in the height of a trace would be a good indicator of high suspense, because the suspense would provoke erratic operation of the sensor. Conversely, it could equally be argued that a low variation was a better indicator of high suspense, because participants feeling suspense could be gripped by the experience to the extent that they did not concentrate on operating the sensor at all. Either way, analysis of amount of change in pressure level and its relationship to response time would be revealing. Therefore, the dataset was subjected to such an analysis.

Several procedures could potentially be used to compute downsampled values representing amount of change in pressure variation from the original data. One method would be to calculate the *first derivative* of the original trace, to convert each point into an *absolute* (i.e. unsigned) value and then downsample the result as described earlier; another, equivalent approach would be to downsample the original trace but to calculate the *variance* under each triangle rather than the mean. The data preparation program (see Appendix A) was adapted to perform the second of these procedures, resulting in a dataset consisting of 266 rows, each of which contained 5 samples representing the amount of change in pressure level over approximately a 10-second segment of a trial in the first part of the experiment, and the associated response time from the second part of the experiment. These data were subjected to a stepwise Multiple Linear Regression analysis, using the same cutoff points as earlier. Independent variables were SAMPLE1, SAMPLE2, SAMPLE3, SAMPLE4 and SAMPLE5; the dependent was RESPTIME.

The analysis yielded two significant models. The first involved only SAMPLE3; the second involved both SAMPLE3 and SAMPLE4. An ANOVA calculation indicated that the first model was significant ($F_{(1,263)} = 5.06$; $p < 0.025$); it could account for 1.5% of the variance in RESPTIME ($Multiple\,R^2 = 0.019$; $Adjusted\,R^2 = 0.015$); the beta-coeffient of SAMPLE3 was -0.14. In other words, response time increased as variation in pressure level about half way through each pressure trace decreased. The second model was highly significant ($F_{(2,262)} = 4.81$; $p < 0.009$) and could account for 2.8% of the variance in RESPTIME ($Multiple\,R^2 = 0.035$; $Adjusted\,R^2 = 0.028$). As expected, its major contributor was SAMPLE3 ($\beta = -0.21$; $t = -0.30$; $p < 0.003$); however, SAMPLE4 also made a significant contribution ($\beta = 0.15$; $t = 2.12$; $p < 0.035$). This could be interpreted to mean that response time increased as the variation in pressure level about half way through each extract decreased, but also increased as the variation *increased* further toward the end of an extract. An alternative and slightly more comprehensible description is that response time lengthened as pressure level variance increased from a low-point in the middle toward the end of each trace. Whichever description is chosen, the results defy obvious explanation. We have a significant result, but it relates in no simple way to either of our hypotheses concerning the possible relationship between variation in pressure level and response time.[12]

It should be noted that whilst the second regression model is highly statistically significant, it cannot account for much of the variance in response time, a state of affairs which is perhaps unsurprising given the vast number of uncontrolled factors that an experiment such as this inevitably involves. When being debriefed after experiment sessions, several of the participants expressed surprise at the notion that response times in the second part of the experiment could be affected in any way by the experience of participating in the first part. Perhaps these participants were right. It is quite possible that response times were in essence an "independent" variable, dictated solely by knowledge and perhaps even emotional salience of the facts; variation in pressure level could be affected by this same factor.

This final suggestion allows us to promote one tentative explanation for the fact that a significant regression was found relating pressure sensor data to response times, and was not congruent with our original hypotheses. Perhaps the profile of the pressure traces varied not so much as a function of perceived suspense but as a function of arousal in response to the music-proposition complex. For example, better known or more salient facts could have evoked shorter response times and also led to high arousal, reflected indirectly in the pressure traces.

Rating Scales

After each trial in the first part of the experiment, participants were asked to rate on a seven-point scale how well the musical excerpt suited the proposition; these ratings were effectively a measure of the extent to which participants felt the musical excerpts and propositions about well-known historical facts to be integrated into a single experience. As such, analysis of them might help shed light on the pressure sensor data reported above: if the profile of the pressure traces was related to arousal as has been suggested, we might expect that trials rated as more highly integrated would provoke traces with a higher pressure level and higher variation in this level. Such a finding would be congruent with an observation reported in Chapter Seven, namely that highly moving music-story complexes led to more cross-domain integration.

In order to test this hypothesis, the seven-point ratings were normalized per participant

[12] Further analyses of the traces using different numbers of downsampled values has failed to shed any more light on the issue.

and paired with their associated pressure traces. From each trace, a mean pressure value was calculated using the triangle-function downsampling procedure with a single triangle covering the entire extent of the trace; this value represented a weighted mean pressure with most of the emphasis placed on the middle of each trace. A two-tailed Pearson's Product Moment Correlation analysis relating mean pressure level with rating revealed a highly significant positive correlation ($r = 0.58$; $p < 0.0001$). In other words, trials rated as highly integrated really did evoke higher pressure-level traces than their less well integrated counterparts. An equivalent analysis comparing rating with mean variation in pressure trace also revealed a highly significant positive correlation ($r = 0.34$; $p < 0.0001$); trials rated as highly integrated evoked traces whose pressure level varied significantly more than their less well integrated counterparts. Finally, a highly significant correlation was found between a downsampled first derivative of the variance and rating ($r = 0.31$; $p < 0.0001$). These results suggest that the overall characteristics of the pressure traces may well have been a function of arousal rather than suspense specifically. This notion in turn lends credence to the proposition that the significant regression found earlier was indeed related to response time being an indicator of emotional salience of the events and the pattern of the trace being at least partly determined by arousal whose extent varied as a function of the music excerpts' suitability as a carrier.

If both short response time and high integration were a function of emotional salience, we should expect to find a negative correlation between rating and response time. This was indeed the case: a one-tailed Pearson's Product Moment revealed a significant correlation ($r = -0.13$; $p < 0.03$).

General Discussion

This experiment certainly did not offer unambiguous results with respect to the initial hypothesis concerning the relationship between anomalous suspense and factual recall. If anything, analyses of the pressure sensor data and response times threw up more questions about the use of pressure sensor data than answers about narrative! In conjunction with the experiment reported in Chapter Seven, however, it does point to a robust link between movingness or arousal and narrative integration; listeners readily bind context into the dynamics of a coherent experience.

Analysis of the results highlighted several practical and methodological difficulties that may well have impacted on the results; a future replication of this experiment that attempts to correct some of these may well find significantly less noisy data and be able to present clearer findings. The most obvious problem concerned the "well-known" facts, which were clearly not very well known by many of the participants. It has been noted that on average, participants correctly judged just more than 15 out of 18 facts correctly; this is a success rate of only 86% in an experiment whose premises relied on a 100% rate. Even this figure of 86% correct is probably an optimistic measure of the number of facts participants actually knew: as has already been noted, the basic laws of probability suggest that if the average participant incorrectly judged 14% of the facts, he or she was actually ignorant of 28% but just happened to judge some correctly by chance. A future experiment should match participants to sets of events more carefully!

A second potential problem was that the musical extracts were relatively similar to each other; all were excerpts from orchestral film soundtracks of a similar genre. This choice of excerpt was made quite deliberately in order to ensure plausibility in the pairing of excerpts and propositions and to provide a genuine possibility that each extract had the potential to evoke suspense for the participants. In retrospect, however, the ends of the experiment may have been better served by more varied musical excerpts because the ones presented here may have been too interchangeable, leading to a range of reactions from participants—and therefore pressure traces for analysis—that may have been somewhat narrower than it might otherwise have been. Corroboration for this possibility comes from a note written in rating-sheet margin by one participant: "you've already played me this one".[13]

Perhaps the most serious difficulties concerned the pressure sensor, both in terms of its use from participants' perspective and in terms of what its data represents. As has already been suggested, it appears that participants may have found the task asked of them too difficult, leading to pressure traces whose height at any given time was not a true reflection of suspense. The combination of a significant regression of downsampled variance and the negative correlation between downsampled mean and rating scale suggest that, if any-

[13]Having seen this comment, the randomisation algorithm in the computer program responsible for ordering and presenting trials was audited thoroughly and found not to be at fault; the participant was incorrect in his diagnosis.

thing, pressure level was more a reflection of arousal than consciously perceived suspense (although the two are surely related in some way). An on-line measuring device, such as a pressure sensor, remains a potentially invaluable way of collecting data which does not rely on retrospective introspection; the very fact that the significant regression found here concerned a dimension that could not plausibly have been plotted consciously by participants is a good indication of the potential of such devices. Much more methodological research is needed, however, if we are to assess exactly where and how pressure sensors can usefully be employed in experiments such as this. As has been noted both at the start of this chapter and elsewhere in the thesis, online measuring devices have become popular recently as a tool for music psychology research, especially in the field of emotion where pressure traces or equivalents are often used to correlate real-time reported emotion with musical features. As Schubert (1998, 2000) has recently demonstrated, however, statistical analyses in such experiments have often been problematic. The results of this experiment suggest that the interaction between participant and device may also be in need of scrutiny.

Despite all the difficulties, the results of this experiment do point to an interaction—albeit a complex one—between the experience of listening to music as reflected in a pressure trace and the presentation of events concerning antecedents of well-known facts. As has been noted, context provided by a statically presented proposition affected the *dynamic* of a listening experience. This fact alone is surely evidence for the involvement of some form of narrative processing in affective response to music. At the very least, it supports the claim that a full understanding of emotional responses to music must consider narrative contexts to which music so readily binds and within which it is always heard. Such an understanding may emerge from research methodologies whose inspiration comes from fields beyond music psychology itself.

Chapter 9

Arousal, Integration and Narrative Process

Despite the growing body of circumstantial evidence pointing to the likely involvement of narrative processes in emotional response to music, one important question remains wide open: can we find a methodology within which narrative processes hypothesized to be implicated in the evocation of emotional response to music can be reliably detected and observed? The two experiments presented so far have investigated aspects of narrative by considering the relationship between music listening and other domains experience of which involves narrative processes. Both of these experiments have revealed interesting findings, but neither has unambiguously answered the methodological question posed here (the first experiment, of course, did not set out to do so). In keeping with one of the goals of the second part of this thesis, namely to illustrate preliminary empirical approaches to the investigation of narrative processes and their implication in emotional response to music, this chapter will present one final experiment that draws on the results of the previous two and a somewhat different stance toward non-musical domains in an attempt to evidence directly the operation of narrative processes in the experience of listening to music.

Both of the experiments already presented demonstrate the extent to which listeners bind the experience of listening to music with stimuli from other domains; this finding is corroborated by previous empirical work that has shown the extent to which music can affect interpretation of a film scene (see e.g. Bullerjahn and Gueldenring, 1994; Vitouch, 2000, and Chapter Five of this thesis).[1] If our interpretation of the pressure sensor data from the most recent experiment is correct, then we also have strong evidence that level of integration between music and surrounding context increases as a function of the extent to which a listener is moved. These two findings suggest a potential approach to the problem of evidencing narrative processing: if narrative processes are involved in the experience of listening to music, and if music listening is integrated with other domains, then exposure to a non-musical stimulus in the presence of music should cause that stimulus to be experienced, at least in part, in a narrative mode. Moreover, the greater the integration between the experience of listening to music and the experience of the non-musical stimulus, the more likely it is that the non-musical stimulus will be appraised as narrative; in the light of the Chapter Eight experiment, we can extend this hypothesis to claim that narrative appraisal of a non-musical stimulus should increase as a function of the extent to which a listener is moved by the experience. In short, the involvement of narrative processes in the experience of listening to music can be evidenced by the presence of narrative processing of some non-musical stimulus; the problem of detecting and observing narrative in music becomes a problem of detecting and observing narrative in non-music. This observation constitutes the basis of the experiment reported here.

In this experiment, participants were presented with a series of photographs each of which depicted either human activity, the results of that activity or a human situation. For example, one photograph showed a balloon salesman walking through a park, another the launch of a shuttle, and another an elderly man sitting by a hospital bedside. As photographs are static images, they are not *narrative* in any literal sense; they cannot tell a story or directly depict action.[2] On the other hand, they are able to capture scenes from which an active mind can infer narrative; they can certainly be seen as representing a "cause and effect teleology", to use Branigan's (1992) terminology (see Chapter Five of this thesis). All of the photographs used in the experiment were chosen because of their potential either to be seen merely as a scene ("balloon salesman in a park") or to be appraised in a narrative mode ("a balloon salesman walking through a park").

[1] One important difference between the experiments presented in this thesis and the previous empirical work mentioned here is that the previous work is concerned only with the effect of music *on* film (i.e. music as context for film) whereas experiments in this thesis are concerned with the experience of listening to *music*; multimedia contexts are only of interest to us here to the extent that they can inform us about emotional response to music.

[2] Photographs that employ special effects such as a long shutter exposure, deliberately to give a sense of motion, are perhaps excepted from this generalisation.

After looking at a photograph, participants were shown four proposed titles for it: two of these titles encapsulated the detail of the depicted scene, whereas the other two emphasized a human, teleological, narrative interpretation. The task was to rate the appropriateness of each of the titles to the photograph. Likewise, participants were shown a list of twelve descriptive words, six of which suggested teleology and six of which did not; the appropriateness of each word was also to be judged. If the act of scrutinizing the photographs employed narrative processes, participants should have been biased towards a narrative interpretation of the scene, thus leading to higher relative ratings for those titles and words that were better matched to narrative interpretation.

For half of the participants in the study, the stimuli were not just photographs, but photographs presented in conjunction with music. Each photograph was paired with a musical excerpt chosen at random from a list that covered a range of moods and styles. While the music played, these participants were asked to provide a continuous measurement of the extent to which they were moved by the listening and viewing experience by squeezing a pressure sensor; afterwards, they too were asked to rate the suitability of titles and words to the photographs. In addition, they were asked to rate the extent to which the listening to the music and looking at the photograph were bound into a single integrated experience. If listening to music involves narrative processes, then participants who listened to music should have been more likely to give high ratings to titles and words that suggested narrative that those that did not; furthermore, the more integrated the experience of listening to the music and observing the photograph, the higher the ratings for narrative titles and words with respect to non-narrative ones. Finally, the more a participant was moved by a trial, the higher still should be the ratings for narrative titles and words with respect to non-narrative ones. This last hypothesis relies on pressure sensor data from the Chapter Eight experiment; however, the correlation between pressure level and integration is tested again here thanks to the inclusion of an integration rating.

Given the steadfast avoidance of rating scales in the Chapter Eight experiment, their use here may seem to be something of a backward step. It should be noted, however, that participants were not asked to use the scales to make introspective judgements (except when assessing integration). Neither were they asked to use them to rate any aspect of *music* or their response to it. Instead, the scales were used to collect participants' judgement of the suitability of title and words to a photograph, a task quite oblique from judgement of a narrative listening process that the experiment set out to detect. Therefore, the experiment is absolutely in keeping with the goals set out in Chapter Eight, namely to avoid retrospective introspection of emotional response to music; it is also in keeping with a core theme of the thesis, namely to consider emotional response to music from the perspective of a listener, in a rich multimedia context.

Method

Participants were presented with a series of 15 photographs. For each photograph, the task was to rate the appropriateness of 4 proposed titles and 12 words for describing the depicted scene. Half of these words and titles suggested a narrative interpretation, whereas the other half did not. One group of participants heard a musical excerpt accompanying each photograph; in addition to the task described above, these participants were asked to squeeze a pressure sensor to reflect the extent to which they were moved by each trial; after rating the titles and words, they were also asked to rate the extent to which music and photograph were integrated into a coherent experience.

Stimuli: photographs

All photographs were taken from the *International Affective Picture System (IAPS)* (Center for the Study of Emotion and Attention [CSEA-NIMH], 1999), a collection of approximately 700 digitized photographs all of which have been rated for affect using empirical methods (see Lang, Bradley, and Cuthbert, 1999); IAPS photographs are mostly of real life scenes, and cover a vast range of subjects and affects. Use of IAPS ratings provided an independent measure of the potential emotion-evoking properties of the photographs.

Photographs were chosen from the collection according to three criteria: each must involve a human element (human activity, the result of human activity, or a human situation); each must be rich enough to withstand viewing for between 30 and 45 seconds; inasmuch as it is possible, each must have an equal potential to be appreciated in a narrative or non-narrative mode. In addition, the set of photographs were selected to cover a wide affective range: specifically, 5 photographs were chosen to represent each of 3 groups: highly arousing and negatively valenced; highly arousing and positively valenced; neutral. Although this categorisation was not essential to the main experimental hypothesis, it was felt that distinguishing the set of photographs in this way may yield some useful additional findings.

To qualify as highly moving and negatively valenced, a photograph had to have a mean IAPS arousal rating of at least 4.5 (out of a possible 9) and a mean valence rating lower than 3.0 (also out of a possible 9). Highly moving and positively valenced photographs also had to have an arousal rating of at least 4.5, and a valence rating of more than 6.0). Neutral photographs had an arousal rating lower than 4.0 and could have any valence rating.[3]

A brief description of each chosen photograph, along with its mean IAPS valence and arousal ratings, are listed in the tables below. Also included in the table is the IAPS number, a 4-digit number that uniquely identifies the photograph within the IAPS collection.

Highly moving—negatively valenced

Description	*Valence*	*Arousal*	*IAPS No.*
Old man by hospital bedside	1.95	4.53	2205
Child running from soldier	2.19	6.01	6212
Passengers from a crashed plane	2.43	6.36	9050
Children playing by powerstation	2.31	5.69	9520
Bomb blast	2.96	6.06	9630

Highly moving—positively valenced

Description	*Valence*	*Arousal*	*IAPS No.*
Space shuttle launch	7.01	5.84	5450
Diver swimming with shark	6.33	5.34	5622
Climber on mountain ridge	7.03	6.55	5629
Table laid for dinner in a garden	6.65	4.89	5849
Man with his two grandchildren	8.03	4.90	2340

[3]Although an arousal rating of 4.5 out of 9 may seem like a rather low threshold for a photograph to qualify as highly moving, it should be noted that the IAPS collection contains a number of photographs of an extremely disturbing or provocative nature; as the experimenter had no wish either to offend participants or to undergo scrutiny by an Ethics committee, no such photographs were included in the experiment set.

Neutral

Description	*Valence*	*Arousal*	*IAPS No.*
Judge sitting in court	4.39	3.07	2221
Holiday couple waving by the sea	6.89	3.09	2501
Balloon salesman in a park	6.64	3.83	2791
Man with bicycle on a hillside	6.03	3.29	5875
An office store cupboard	4.25	2.95	7708

Each of these photographs was already in JPEG format at a resolution of 800x600 pixels, and could therefore be loaded directly into a computer without manipulation, ready for full-screen display by the experiment control software. Photographs were presented to participants in a random order; the experiment control software generated a new random sequence for each participant. For participants in the NO-MUSIC group (see below), each photograph was displayed for 30 seconds; for participants in the MUSIC group, the photograph remained on screen for the duration of its accompanying musical excerpt.

Stimuli: titles

For each photograph, 4 titles were created. Each of the titles in a set constituted appropriate descriptions of the depicted scene, but they differed according to two categorical dimensions: NARRATIVITY and MOVINGNESS. In terms of the main experimental hypothesis, the more important dimension was NARRATIVITY: two of the titles suggested a narrative interpretation of the photograph, whereas the other two did not. Within each of these categories, the MOVINGNESS dimension distinguished one title which used plain, factual language from the other, which used more emotionally suggestive language. This second variable was included to preclude the possibility that a difference of ratings between narrative and non-narrative would be due to the narrative titles being intrinsically more arousing than the non-narrative ones.

Unmoving, non-narrative titles were created by describing the core depiction of each photograph as concisely as possible, whilst avoiding descriptions that suggested teleology or agency. The other three titles were generated from this base according to the following formula: narrative titles were created by changing the wording to emphasize teleology or action; although the exact nature of the change varied idiosyncratically to suit the specifics of each photograph, this usually meant inserting or replacing a gerund with an active verb. Moving titles were formed from the plain bases either by inserting emotive adjectives, or by replacing words with more emotive equivalents.

The title sets for each of the photographs are listed below. The order in which the sets are presented here matches the order of the photographs as listed in the tables above; within each set, the order is as follows: plain/non-moving, plain/moving, narrative/non-moving, narrative/moving.

an old man by his ill wife's bedside
a frail old man by his ill wife's hospital bedside
an old man sits by his dying wife's bedside
a frail old man sits by his dying wife's hospital bedside

child in line of soldier's gun
helpless child in line of soldier's lethal gun
child runs from soldier's gun
helpless child runs from soldier's lethal gun

passengers in front of burning plane
shattered passengers in front of fiercely burning plane
passengers escape as plane burns
shattered passengers escape as plane burns fiercely

children playing in polluted water near powerstation
small children playing in filthy polluted water near powerstation
children play as nearby powerstation pollutes their air
small children play as nearby powerstation poisons their air

bomb blast in night sky
searing bomb blast in dark night sky
a bomb blast illuminates the night sky
a searing bomb blast illuminates the dark night sky

space shuttle at takeoff
fire-filled sky at space shuttle takeoff

space shuttle embarks on its voyage

fire-filled sky as space shuttle embarks on its voyage

diver with a shark

diver with a large shark

diver swimming with a shark

diver swimming with a large shark

climber on mountain ridge

climber on precarious mountain ridge

a climber traverses mountain ridge

a climber traverses precarious mountain ridge

dinner table in the garden

lavish dinner table in the garden

waiting for dinner in the garden

waiting for a lavish dinner in the garden

a [joyfully] smiling grandfather with grandchildren in his arms[4]

a smiling grandfather with happy grandchildren in his arms

a grandfather smiles as his grandchildren play in his arms

a grandfather smiles joyfully as his grandchildren play happily in his arms

a judge in court

a severe judge in court

a judge brings court to order

a severe judge brings court to order

a holiday couple in front of the camera

a happy holiday couple in front of the camera

a holiday couple pose for the camera

a happy holiday couple pose for a photo

balloon man with his balloons in a park

balloon man with his bright cluster of balloons in a park

[4]The inclusion of the word "joyfully" in this plain, non-narrative title was an error that went undetected until several participants had already completed the experiment. It was decided, therefore, to leave the word in place. As an analysis contrasting moving and non-moving titles would involve 30 data points per participant, leaving this anomaly would have no significant effect on results.

balloon man walking his balloons through a park

balloon man walking his bright cluster of balloons through a park

man with a bike on a hillside

old man with a bike on a golden autumnal hillside

man pauses for breath while pushing his bike up a hillside

old man pauses for breath while pushing his bike up a golden autumnal hillside

store containing old documents

store containing dank old documents

stowing away old documents

stowing away dank old documents

In a pilot study, a small group of people who did not participate in the main experiment were given the complete list of titles as presented above and asked to identify the category to which each belonged and the photograph to which each referred. This task was accomplished with a 100% success rate. All titles were typed into a text file in a format suitable for being read by the experiment controlling program. The program was responsible for ensuring that the correct set of titles was presented to each participant after the display of each photograph; within each set, titles were presented in an order determined at random for each participant.

Stimuli: words

A set of 12 descriptive words were selected, 6 of which were intended to suggest teleology or agency whereas 6 were not. Within each of category, half of the words were positively valenced and half negatively valenced; all were intended to be reasonably evocative. A combination of polarised valence and high arousal meant that half of the words were likely to be highly unsuitable as descriptions for any given picture; this was not considered to be a problem, however, because the same set of words was used for all trials, allowing direct comparison of mean scores within each category.

All words were taken from the *Affective Norms for English Words* database (Bradley and Lang, 1999), a database of approximately 1000 words all of which have been affectively rated (using the same scales as the IAPS photographs). Words were selected according to the following criteria:

- Each should have a mean ANEW arousal rating of at least 7.0.
- Positively valenced words should have a mean valence of at least 6.5, and negatively valenced words should have a mean valence of less than 3.0.
- Each should be generic enough to be applicable to a wide range of situations, so that each word was likely to be deemed suitable for several of the photographs.
- Half of the words must imply agency or teleology; the other half must not.

These criteria left the experimenter with a very small pool from which the final selection was made. The chosen words, together with their ANEW ratings and identification numbers, are presented in the tables below.

Narrative words

Word	*Valence*	*Arousal*	*ANEW No.*
terrified	1.72	7.86	432
danger	2.95	7.32	713
trouble	3.03	6.85	454
excitement	7.50	7.67	152
triumphant	8.82	6.78	452
adventure	7.60	6.98	630

Non-narrative words

Word	*Valence*	*Arousal*	*ANEW No.*
brutal	2.80	6.60	53
horror	2.76	7.21	213
violent	2.29	6.89	478
fun	8.37	7.22	759
festive	7.30	6.58	749
joy	8.60	7.22	240

All of the words were typed into a text file in a form suitable for reading by the experiment control program, which was responsible for randomizing their order prior to each trial for each participant.

Stimuli: music

The hypotheses that this experiment attempted to test did not rely on musical excerpts fulfilling any specific criteria; it was only important that the various combinations of musical excerpts and photographs with which participants would be presented would be likely to evoke a range of emotional responses. To that end, a selection of 15 musical excerpts was chosen—one for each trial—that covered a range of different styles and genres. This music included six excerpts from the experiment reported in Chapter Eight; in addition, there were extracts from two romantic film soundtracks, three classical orchestral and piano works, one contemporary classical work, one pop song and two rock songs. Full details of the pieces and discs from which they were taken can be found in Appendix B; the excerpts themselves are reproduced on *Compact Disc Two - Stimuli.*

It was decided that each excerpt should last for between 30 and 45 seconds, as this was deemed to be approximately the amount of time that a photograph could reasonably be displayed for participants who would not be presented with a musical accompaniment. Therefore, a suitable segment of music from each of the original recordings was chosen and then prepared by loading the relevant CD tracks into a computer-based digital editor and applying smooth fades to the ends (and, where necessary, the starts). All 15 edited stimuli were then converted to MP3 format (using the *xingtech* variadic bit-rate encoder) and loaded on to the experiment-running computer.

Participants heard each of the 15 excerpts only once; they were randomly ordered, so that every participant experienced a different combination of music and photographs.

Apparatus

Experiment sessions were controlled by a PC-based Linux workstation running software written specifically for the purpose (see Appendix A). This software was responsible for creating random orderings of photographs, titles, words and music, presenting the stimuli and collecting responses. Sound and video hardware was the same as that used in the experiment reported in Chapter Eight. Photographs were displayed in a high-resolution "full-screen" mode so that whilst they were presented, nothing else cluttered the screen. The pressure sensor used to measure the extent to which participants were moved by each trial was the one designed and built for the experiment reported in Chapter Eight.

Participants

24 people participated in this experiment all of whom were students or staff at Cambridge University. 4 participants were studying music; none studied psychology. 12 participants were assigned to the MUSIC group (i.e. were presented with photographs accompanied by the musical excerpts) and 12 to the NO-MUSIC group (i.e. were presented with photographs but no music; did not use the pressure sensor or rate level of integration). Assignation of a particular participant to the MUSIC or NO-MUSIC group was arbitrary, based on the experimenters continuous assessment of how many people were likely to participate in total, so that he could ensure that an equal number of participants would have been assigned to each group once the sessions were all complete.

Procedure

Participants took part in the experiment individually. On arrival in the experiment room, those who had been assigned to the NO-MUSIC group were asked to sit in front of the computer and were presented with the following on-screen instructions:

> *Please read these instructions carefully. If you are at all unsure of any points, do not hesitate to ask the experimenter. In this experiment, you will be presented with a series of fifteen trials, each of which will take the following form:*
>
> *A photograph will appear on the computer screen depicting a real-life scene. Look at this photograph and reflect upon the scene it depicts. You will not be asked to recall any of its details later on, so there is no need to scrutinise it minutely as if for a test; rather, try to relax and reflect upon the scene naturally as if you are seeing the photograph in a newspaper or on television.*
>
> *After approximately 30 seconds, the photograph will disappear and be replaced with a window containing two distinct areas, corresponding to the two tasks you are asked to*

> *perform. The top area will list four proposed "titles" for the photograph that you have just seen. Next to each title is a seven-position slider. Your task is to rate how well you feel each title captures or reflects the scene depicted in the photograph. If you feel that a title captures the photograph extremely well, you should move its companion slider to the far right-hand position; if you feel that it captures it just about satisfactorily, leave the slider in the far left-hand position. The other five positions should be used to represent ratings between these two poles. All titles presented will be appropriate to the photograph, so you should not be restricted by a scale whose minimum rating is "just about satisfactory". Please note that you are NOT being asked to assess the artistic or journalistic merits of the titles* qua *titles; rather, you should judge subjectively how well you feel each one reflects or captures the photograph, using as your guide your thoughts and feelings as you saw it.*
>
> *The second area contains a list of words; next to each word is a seven-position slider. Your task here is to judge how appropriate you feel each word is for the scene depicted in the photograph that you have seen. If you feel a word to be extremely appropriate, move its companion slider to the far right-hand position. If you feel that it is completely inappropriate, leave the slider in the far left-hand position. The other five positions should be used to represent ratings between these two poles.*
>
> *Once you have finished rating the titles and the words, click the button at the bottom of the screen and the next photograph will be displayed. If you would like a short break between any trials, just abstain from clicking the button until you are ready to move on.*

Having read the instructions, the experimenter confirmed that participants understood the task asked of them and then allowed the trials to run. The experimenter remained in the room for the duration of the first trial to ensure that participants were comfortable using the software; he then left participants to complete the remainder of the trials without distraction. An experiment session lasted for approximately 20 minutes.

Procedure for participants assigned to the MUSIC group was similar to that for the NO-MUSIC group, but with a few important differences. The second paragraph of the experiment instructions was modified to read as follows:

> *A photograph will appear on the computer screen depicting a real-life scene; it will be*

accompanied by a musical excerpt. Look at this photograph and reflect upon the scene it depicts. You will not be asked to recall any of its details later on, so there is no need to scrutinise it minutely as if for a test; rather, try to relax and reflect upon the scene naturally as if you are seeing the photograph in a newspaper or on television. To your right, you will see a pressure sensor; as you reflect upon the photo and the music plays, use the pressure sensor to record your the extent to which you are moved by the experience of looking at the photograph and listening to the music.

Two other more minor modifications to the instructions reflected the fact that each photograph would appear for a length of time determined by the duration of the musical excerpt (approximately between 30 and 45 seconds) and that participants would be asked to rate the extent to which listening to the music and looking at the photograph constituted an integrated experience.

Unlike in the experiment reported in Chapter Eight, participants here did not have visual feedback from the pressure sensor because this would have interfered with the display of the photographs. Therefore, between reading the instructions and starting the experiment trials, participants in the MUSIC group were invited to spend a short amount of time exploring the sensor's range and behaviour, using a visual dial that the experiment-control software displayed for the purpose. Having familiarised themselves with the sensor, participants were invited to put on the headphones and the first photograph was displayed. After the first trial had been completed successfully, the experimenter left the room. An experiment session lasted for approximately 25 minutes.

Design

Test of the first hypothesis—participants in the MUSIC group would give higher ratings to words and titles compatible with a narrative interpretation of the photograph than to those that were not; no such effect would be found for participants in the NO-MUSIC group—consisted of analyses of variance involving two dependent measures: TSCORE (appropriateness rating given to a title) and WSCORE (appropriateness rating given to a word). Conditions for TSCORE measures were manipulated by one between-group variable, GROUP, which specified whether or not a participant were presented with the photographs in the presence of music (MUSIC or NO-MUSIC), and three within-group variables: PTYPE, NARRATIVITY, and MOVINGNESS. PTYPE specified whether a photograph was highly-moving/negatively valenced (HN), highly-moving/positively valenced (HP), or neutral (N); NARRATIVITY specified whether a title suggested narrative (N) or was a non-narrative description of the scene (P); MOVINGNESS specified whether a title was phrased to be moving (M) or non-moving (NM). WSCORE conditions were also manipulated by GROUP and PTYPE, but NARRATIVITY and MOVINGNESS were replaced with WNARRAT, which distinguished narrative (N) from non-narrative words (NM).

Test of the second hypothesis—the greater the integration between the experience of listening to the music and reflecting upon the photographs, the greater the preference for narrative words and titles over their non-narrative counterparts—was restricted to data from participants in the MUSIC group. It consisted of a number of correlation analyses involving two measures—INTEGRATION and PRESSURE—in conjunction with transformations of TSCORE and WSCORE. For each trial, PRESSURE represented the weighted mean pressure trace as calculated by the downsampling procedure described in Chapter Eight, and INTEGRATION represented the participant's assessment of the extent to which the experience of listening to the music and looking at the photograph were integrated into a coherent whole.

Results and Discussion

The presentation of results is split into two distinct sections. First, the first hypothesis will be considered with an analyses of TSCORE and WSCORE with respect to PTYPE, NARRATIVITY, MOVINGNESS and GROUP. Then, the second hypothesis will be considered in the light of an examination of INTEGRATION and PRESSURE.

Testing the First Hypothesis

The first test consisted of a three-within, one-between repeated-measures analysis of variance on the dependent TSCORE. Within-group variables were PTYPE (3 levels), NARRATIVITY (2 levels) and MOVINGNESS (2 levels); the between-group variable was GROUP (2 levels). There were no significant main-effects of any of the within-group variables. There was, however, a significant two-way interaction between PTYPE and NARRATIVITY ($F_{(2,44)} = 3.19$; $p < 0.05$); this was characterised by a pronounced difference between levels of NARRATIVITY in

PTYPE levels HN and HP but not for level N. Specifically, titles that were suggestive of narrative evoked higher appropriateness ratings than those which did not in the HN and HP conditions. We will return to this finding below. MOVINGNESS was neither a main effect nor was it involved in any interactions.

An overall significant effect of GROUP ($F_{(1,22)} = 6.13$; $p < 0.02$) was caused by participants in the NO-MUSIC group giving higher appropriateness ratings than those in the MUSIC group. Although not predicted, this result was not entirely surprising: the titles were all created to match the photographs to which they referred; unless all of the musical excerpts were always an excellent match for the photographs, it should be expected that the set of titles would be less appropriate to the music-photograph complex than to the photographs when presented alone.

According to the first hypothesis, we should have expected a significant two-way interaction between GROUP and NARRATIVITY, but none was found; GROUP was not involved in any significant interactions at all. In other words, the presence of music did not increase participants' preferences for titles that suggested narrative. On the basis of this result, we cannot support the first hypothesis.

A two-within, one-between repeated-measures analysis of variance of the dependent WSCORE yielded similar findings. Within-group variables for this analysis were PTYPE (3 levels) and WNARRAT (2 levels); the between-group variable was once again GROUP (2 levels). Here, there was a highly significant main effect of PTYPE ($F_{(2,44)} = 67.85$; $p < 0.0001$) and of WNARRAT ($F_{(1,22)} = 29.70$; $p < 0.0001$). The effect of PTYPE was caused by participants giving higher appropriateness ratings to words when presented with the more moving photographs (those representing the HN and HP levels of PTYPE) than the neutral ones (those representing the N level); this finding reflects the fact that all of the words were highly arousing (according to their ANEW ratings) and were probably rather unsuited to the neutral photographs. The effect of WNARRAT was characterised by higher scores being given to the narrative words regardless of other variables. A highly significant two-way interaction between PTYPE and WNARRAT ($F_{(2,44)} = 28.48$; $p < 0.0001$) mirrored that found in the TSCORE analysis: HN and HP levels of PTYPE evoked higher scores for words that suggested narrative than for those that did not.

In this analysis, GROUP did not have a significant main effect, nor did it feature in any significant interactions. Although a two-way interaction between GROUP and WNARRAT ($F_{(1,22)} = 2.59$; $p < 0.12$) came tantalisingly close to marginal significance, this result cannot be used to reject the null hypothesis!

Testing the Second Hypothesis

The second hypothesis stated that the more integrated the experience of listening to music and looking at the photograph, the greater the preference should be for narrative titles and words. As this hypothesis involved analysis of pairs of dependent variables, it was tested with a series of Pearson's Product-Moment correlations, each involving 180 data points (15 trials * 12 participants). All tests were performed using the more conservative two-tailed probabilities; analyses used data exclusively from participants in the MUSIC group.

Two new variables were constructed for these tests: TDIFF and WDIFF. TDIFF was created from each of the four TSCORE measures for each trial according to the following formula: $tdiff = \frac{(tscore[n_1]+tscore[n_2])-(tscore[p_1]+tscore[p_2])}{2}$. In other words, the TDIFF measure for each trial represented the difference between the mean rating for the two narrative titles and the mean rating for the two non-narrative titles; the higher TDIFF, the higher the rating given to the narrative titles with respect to the non-narrative ones. The WDIFF measure was the equivalent for the words. For each trial it was created from WSCORE according to the following formula: $wdiff = \frac{(\sum_{i=1}^{6} wscore[n_i])-(\sum_{i=1}^{6} wscore[p_i])}{6}$, or the difference between mean score for the narrative words and those for the non-narrative words; the higher WDIFF, the higher the rating given to narrative words with respect to non-narrative ones.

Both TDIFF and WDIFF were found to be significantly positively correlated with INTEGRATION (the extent to which listening to the music and contemplating the music were judged by participants to be a single coherent experience; $r = 0.17$; $p < 0.05$ and $r = 0.21$; $p < 0.01$ respectively). The higher the level of integration between music and picture, the higher the appropriateness rating given to narrative over non-narrative titles and words. The second hypothesis is therefore confirmed. That the correlation between WDIFF and INTEGRATION is stronger than the one between TDIFF and INTEGRATION is not entirely surprising, as the words—being more generic than the titles—were probably more susceptible to the influence of interactions between picture and music; also, given that the same set of words was used in every trial, participants had the opportunity to hone their discriminatory skills with respect to them.

An extension to the hypothesis confirmed here is that the more moving the experience, the greater the preference for narrative titles and words. This suggestion was based on a finding of the experiment reported in Chapter Eight, namely that the extent to which music and non-musical stimuli are integrated into a coherent listening experience varies as a function of the extent to which a listener is moved by the experience. Hence, the next analysis attempts to corroborate this finding. For each trial, a PRESSURE value was calculated by downsampling the complete pressure trace to a single point using the algorithm described in Chapter Eight; the result of this process was to create a series of single values, each of which represented the mean pressure level for a trace, weighted toward the middle of the trial. These weighted means were then subjected to a Pearson's Product-Moment (two-tailed) correlation analysis against INTEGRATION ratings. The result was a significant positive correlation ($r = 0.28$; $p < 0.01$). In other words, the Chapter Eight findings were corroborated; the more moving a listening experience, the more music integrates with context and *vice versa*.

Given that TDIFF and WDIFF are positively correlated with INTEGRATION, and that INTEGRATION is positively correlated with PRESSURE, we can deduce that TDIFF and WDIFF should also be positively correlated with PRESSURE. An attempt to corroborate this with a direct correlation analysis between the variables showed that WDIFF was indeed significantly correlated with PRESSURE ($r = 0.23$; $p < 0.01$), but that TDIFF was not; despite the chain of positive correlations, they were not strong enough between them to reveal a direct correlation between the extent to which participants were moved and the extent to which narrative titles were deemed more appropriate than non-narrative ones. The most likely explanation for the disparity between TDIFF and WDIFF here is that the set of 12 words was identical across all trials, whereas the sentences—by necessity—were different for each trial; participants were able to hone their ability to discriminate between the words more than for the sentences. To put this in a different way, variance of TDIFF ratings was affected by much more "noise" than WDIFF ratings. It should be noted that a similar phenomenon was found above in the test of the first hypothesis: the two-way interaction between PTYPE and NARRATIVITY was considerably more pronounced for WSCORE than for TSCORE. This two-way interaction can now, of course, be understood easily: the more a participant was moved by a trial, the higher the preference for narrative words and titles.

A final set of analyses tested the extent to which music and photograph were responsible for INTEGRATION, movingness ratings (as encoded by PRESSURE), TDIFF and WDIFF scores. Four univariate analyses of variance tested each of these dependent variables against two new variables, MUSIC and PHOTOGRAPH. These new variables each had 15 levels, one for each of the musical extracts and photographs respectively. As expected, INTEGRATION was affected by neither music nor photograph alone, but by a two-way interaction between the two ($F_{(94,180)} = 1.58$; $p < 0.03$). PRESSURE showed a similar pattern of results, albeit only with marginal significance ($F_{(88,168)} = 1.40$; $p < 0.095$); this weaker result for pressure is consistent with the correlation findings reported in the paragraph above. By contrast, no significant main effect or interactions involving MUSIC were found for either TDIFF or WDIFF; in both cases, the only factor of relevance was PHOTOGRAPH, whose effect was marginally significant for TDIFF ($F_{(14,180)} = 1.83$; $p < 0.06$) and highly significant for WDIFF ($F_{(14,180)} = 11.44$; $p < 0.001$). In other words, despite the various relationships between movingness, integration, photograph type and the relative weighting given to narrative titles and words, and despite the influence of music on these, the presence of specific musical extracts does not appear to have had a significant effect on participants' propensity to appraise a simultaneously presented photograph in a narrative mode. This observation is considered further below.

Conclusions

This experiment has demonstrated a relationship between the extent to which a listener is moved and the extent to which a photograph is appraised in a narrative mode. Furthermore, it has corroborated the findings of the experiment reported in Chapter Seven, namely that integration between music and context increases as a function of the extent to which a listener is moved by the experience. Finally, it has confirmed that pressure sensor data can provide a reliable measure of movingness or arousal.

One thing that it has certainly not shown, however, is that listening to music *causes* a simultaneously presented photograph to be appraised in a narrative mode; not only did the final analysis fail to reveal any significant effect of MUSIC on TDIFF or WDIFF, but the very first analysis did not even find a significant difference between results for partipants who were presented with music and those who were not! A possible explanation for this finding—or rather this *lack* of finding—is that the musical extracts were not particularly well-chosen and did not cover a wide enough range to evoke the whole gamut of emotional reactions that music is known to be able to elicit in many listeners; after all, whereas the photographs were selected from a collection whose contents have been independently rated for affective content, musical extracts were the idiosyncratic choice of the experimenter. In other words, the data may have been simply too noisy; another study, involving more carefully chosen music—perhaps in conjunction with a larger number of participants—might have solved the problem.

There is, however, another explanation. It is possible that the experiment failed to show an effect of music on the extent to which a photograph was appraised as narrative because, in reality, music *has no effect* on the extent to which a photograph—or any simultaneously presented stimulus, for that matter—is appraised as narrative. To be sure, there is a relationship between the two, but the causal chain could be quite different from that suggested at the start of this chapter. All of the photographs presented in the experiment reported here contained a human element; though static, they all had the potential to be appraised in a narrative way. One of the most persistent claims of this thesis, and the rationale for the Sound-Utterance-Context-Narrative model, has been that emotional response to music is not a special, unique phenomenon but can best be understood in terms of our propensity to respond emotionally to many aspects of our environment; narrative, in particular, has been touted as a generic process underlying coherence of experience and (and after Ortony et al., 1988) emotional appraisal. Surely, then, it is likely that all participants in the experiment would have appraised the photographs in a narrative way—particularly the emotive photographs—regardless of whether or not they were also presented with music. The difference between participants in the two groups was merely that those participants who heard music would have also listened to the music in a narrative mode; the narrative mode would have been triggered by appraisal of the photographs, and a desire to cohere experience. The more a photograph encouraged narrative thought, the more music and photograph would have been bound into an integrated, coherent experience, and the more moving the listening and viewing experience would have become. In short, perhaps music is heard, at least in part, as narrative because the experinece of listening to music is, at least in part, driven by a *context* within which music is always heard and to which it so very readily binds.

Confirmation of this suggestion, of course, would require further research. Given the findings of all three experiments concerning the integration of music and context, however, it does perhaps offer the most plausible explanation for the results reported here; it is also compatible with the central tenets of the Sound-Utterance-Context-Narrative model. All three experiments reported in this thesis have perhaps thrown more questions than provided answers; results of all have suggested refinements that would probably lead to clearer findings. Furthering this research and implementing these refinements is a task that lies beyond the scope of this thesis. Meanwhile, the experiments as they stand have fulfilled their main objective, namely to demonstrate the potential of a research agenda focussed on investigating the involvement of narrative processes in emotional response to music, to illustrate ways in which these processes might be detected and observed, and to suggest empirical methods within which music can be studied not as an isolated phenomenon, but within a rich, multi-domain context.

Chapter 10

A Framework for Empirical Research

The model presented in this thesis conceptualizes emotional response to music as the result of a listening process that hears music as sound, utterance, context and narrative. Part One of the thesis consisted of a theoretical discussion of the four components of the model; the listener-centric distinctions between these components were rationalized in the context of many existing empirical studies that have investigated aspects of emotional response to music and related phenomena. Emotional response to music itself was situated within the wider context of what we know about emotion from fields beyond music psychology. Part Two presented a preliminary empirical investigation of narrative, the fourth component of the model. It consisted of three experiments whose primary purpose was to demonstrate that despite the dearth of music psychological research on the subject to date, a programme focused on understanding the involvement of narrative processes in music listening could reveal much about the dynamics and coherence that are so often hallmarks of emotive listening experiences. As suggested by the subtitle, however, the thesis did not so much set out to present a theoretical model of emotional response to music; rather, it sought to provide a "framework for empirical research". In other words, it sought to present a model within which empirical observations hitherto seen as isolated phenomena could be understood as part of a coherent whole and upon which further music psychological work on emotional response to music could build. This, the final chapter of the thesis, takes up the mantle dropped at the end of Chapter Six in arguing that the Sound-Utterance-Context-Narrative model constitutes just such a framework, and suggests some avenues for future research.

The argument is founded upon five basic observations concerning the nature of the model: it is theoretically complete, generic, empirically founded, can account for existing research, and leads to the formulation of useful research questions. It is theoretically complete in that a full understanding of how we respond to music as sound, utterance, context and narrative would *by definition* constitute a full understanding of emotional response to music with no phenomena left unexplained.[1] It is generic in that its premises are based on human perceptual, affective and cognitive processes rather than elements of a musical object. Although this thesis has for the most part concentrated on emotional response to western music by western european and american listeners within the context of twentieth-century culture, the model itself stands independent of these; its fundamental tenets are posited to apply equally to *any* listener in *any* musical culture. It is empirically founded in that its tenets and the fundamental distinctions that it draws are based upon empirical data from music psychological studies and from research in other related fields.[2] It is able to accommodate existing empirical work: its adoption would not require a huge re-evaluation of existing data, all of which fit comfortably within its ambit. Perhaps the most important facet of the model, though, is that its adoption leads to a hierarchy of clear research questions that lend themselves to empirical investigation. It is to this notion that we now turn.

Understanding emotional response to music within the Sound-Utterance-Context-Narrative model involves answering just one fundamental question: "*how* or *by what means* does hearing music as sound, utterance, context and narrative evoke emotion?". This question, in turn, leads to a series of far more specific ones that are more obviously open to empirical testing. For example, consideration of how or by what means music heard as sound evokes emotion might lead to the sub-question "what makes a sound arousing?", which in turn would lead to questions such as "do certain acoustic signals consistently evoke differentiated responses in human listeners?", "are some sounds universally more emotionally salient than others?", "what are the brain mechanisms by which auditory stimuli arouse emotion centres?", "by what factors are they mediated?", etc. Similar questions might be asked about

[1]By this token it could be argued that the question "why are we are moved by music?" constitutes a theoretically complete framework—albeit not a model—because a full answer to this would also constitute a full understanding of emotional response to music! There is, however, an important difference: the Sound-Utterance-Context-Narrative model claims theoretical completeness even though it distinguishes four discrete components of a listening experience.

[2]As has previously been noted, the narrative component does not find much support from within music psychology; hence Part Two of the thesis. However, as has also been noted, a wealth of empirical and theoretical research in other fields provides a corroboratory evidence for the category, as does work on musical expectancy and expressive timing deviations, which can be interpreted within a narrative frame.

hearing music as utterance: "are some utterances universally more emotionally salient than others?", "which acoustic correlates of emotionally charged vocalization are detected and responded to most readily by listeners?", "what are the brain mechanisms whereby utterances evoke emotion?", "how are these mechanisms distinct from those responsible for response to music as sound?", etc. Consideration of context might evoke a slightly different set of questions, such as "are memories triggered more readily by certain types of music than others?", "what are the mechanisms by which context mediates mood?", "to what extent does mood affect response to music as sound or utterance?", "to what extent do culture-specific conceptualizations and contexts mediate musically-invoked mood?", etc. As for narrative, one might ask: "how are disparate responses to sound, utterances and context integrated into a coherent narrative?", "what are the processes through which narrative structures can evoke emotion?", "what factors mediate the extent to which elements of a listening experience bind together into a cohesive whole?" , "what are the most salient sources of structure and content in narrative derived from a listening experience?", etc. These questions, of course, represent only a tiny subset of those that could be posed, but they give some idea as to the scope of the field as defined by the Sound-Utterance-Context-Narrative model. Some of the questions listed here have already been subject to empirical investigation either directly or indirectly; some have even been answered to a greater or lesser extent. The first part of this thesis reviews a wealth of empirical research that accomplishes just this. Others, by contrast, particularly those concerning narrative, have neither been answered nor even asked; hence the second part of the thesis. Crucially, all these questions, whether investigated or not, emerge from and lie clearly within the ambit of the Sound-Utterance-Context-Narrative model; answers to them can all form part of a coherent picture of emotional response to music.

Not only does the Sound-Utterance-Context-Narrative model lead to specific lines of questioning, but it also suggests methodological approaches within which such questions may be answered. For example, an understanding of auditorily-evoked arousal (immediate responses to sound) lies clearly in the domain of neuroscientific research; an understanding of response to utterance could benefit from developmental psychological and comparative ecological or biological approaches; context and narrative fall more squarely within the remits of social and cognitive psychology. Perhaps the most striking commonality between these suggestions is the lack of anything exclusively musical about them. The reason for this—as Chapter One of the thesis conjectured and as the experiments presented in Part Two confirmed—is that listeners so readily bind extra-musical stimuli as an integral part of a listening experience that it seems perverse to treat the "music" part separately from its context; a dichotomy between intra- and extra-musical features, or between so-called *intrinsic* and *extrinsic* sources of emotion is of far more salience to theorists than listeners. This is not to contest the notion that musical structure is important for western listeners' emotional responses to the music with which they are familiar; it is merely to suggest that musical structure as a source of emotion is not fundamentally different from any so-called extrinsic source. Empirical research focussed on *emotional response* to music need not necessarily focus on music itself.

Throughout this thesis, narrative has been afforded special status and treated rather more extensively than the other components of the Sound-Utterance-Context-Narrative model. As has previously been noted, the preferential treatment is partly due do the lack of previous music psychological work considering emotional response to music as a phenomenon involving narrative process. However, the difference also owes much to the ontological status of narrative within the model: whereas sound, utterance and context are seen as three distinct components whose emotion-evoking properties are likely to be mediated by distinct mechanisms leading to distinct effects, narrative is an umbrella under which every aspect of musically evoked emotion is subsumed. This distinction has repercussions for the model as a framework: subscribing to the notion that music is heard as sound, utterance and context suggests a particular line of questioning for empirical research; subscribing to the notion that music is heard as narrative not only suggests questions, but also dissolves a whole range of them! For example, questions such as "what causes emotional response to music to be coherent?", "what are the mechanisms by which expectancy violations appears consistently to evoke emotion within western culture?", "how do context and musical structure interact in a listening experience?", and many others all collapse into one single question, namely, "how does narrative experience evoke emotion?". Answering this question—within which the word "music" is so conspicuously absent—is central to the present model of emotional response to music.

Results of the experiments presented in Chapters seven, eight and nine lend substantial support to the notion that research questions concerning coherence, expectancy violation and interaction of context and structure can indeed be collapsed into a single question about narrative. As has already been noted, all three experiments have demonstrated that listeners willingly impose coherence upon stimuli from different domains, binding musical with non-musical stimuli; this phenomenon appears to hold whether the non-musical stimuli in question are presented passively (e.g. a photograph, as demonstrated in chapter nine), require active work on behalf of the listener (e.g. a short story, as demonstrated in chapter seven), or even demand recall of facts blatently unrelated to the music (e.g. the sentences presented in chapter eight). Results of the final experiment, presented in chapter nine, also provide more direct evidence to suggest that listeners really do employ narrative processes when listening to music; by extension, this experiment demonstrates that it is possible to construct laboratory conditions in which narrative processes may be observed at work, without resorting to crude introspective techniques. In short, the experiments reported in Part Two of the thesis appear to confirm that the more satisfactorily we can answer the question "how does narrative experience evoke emotion", the more we will understand about emotional response to music.

Despite the encouraging signs, it should be noted that none of the experiments was without its problems; each employed previously untested methods and, as has already been discussed at length, data were relatively noisy and difficult to interpret. Further research will be required before we are able confidently to embrace the concept of narrative within the core of music psychology, and consider the narrative part of the Sound-Utterance-Context-Narrative model to be as non-contentious as the other three components. One possible approach to this further research would be to refine the methodologies used in the experiments presented here; several possible improvements and alterations were suggested in the experiment reports themselves, ranging from choice of stimuli to type of measurement. Alternatively, completely new experimental techniques could be designed with the aim of observing and perhaps isolating narrative processes during a listening process. From the data already gathered, it seems that development of a multimedia approach, involving the manipulation of music and filmic or pictorial contexts may yield useful results. Either way, before the Sound-Utterance-Narrative model can stand as a fully viable framework for empirical research, a reasonable corpus of experimental work investigating the crucial narrative component will have to be undertaken.

Further research on narrative will not merely strengthen the basis of Sound-Utterance-Narrative-Context as a model; it will also—in conjunction with the model—provide a much-needed link between what we know about music listening and what emotion theorists tell us about emotion. Chapter Six of the thesis situates the Sound-Utterance-Context-Narrative model within a conceptualization of emotion or affective response as a combination of physiological arousal and cognitive explanation of that arousal (which can itself also be a source of arousal cf Schachter and Singer 1962). Sources of arousal are easily identifiable within the terms of the model: they are sounds, utterances, context-dependent memories and the like. By contrast, cognitive explanation of that arousal is not so unambiguously identifiable. To be sure, it may include recognition of the *significance* of a sound, utterance or context with respect to the well-being of the listener, but this observation alone falls a long way short of an explanation compatible with a fully-fledged appraisal theory of emotional antecedents, such as that of Ortony et al. (1988). The missing link is provided once again by the concept of narrative: if "the essence of narrative is a presentation of systematic change through a cause and effect teleology" (Branigan, 1992, p.19), and if such a presentation "encapsulates the interest we take in the world as humans"(Branigan, 1992, p.29) then narrative provides a cognitive framework entirely compatible with that described by appraisal theorists, which sees the cognitive basis of emotion as deriving from "events and their consequences, agents and their actions, or object, pure and simple"(Ortony et al., 1988, p.13). In short, to subscribe to the notion that music is heard as narrative is to subscribe to the notion that when people listen to music, they attempt to make sense of what they hear terms of cause and effect, in terms of a narrative world, and respond to this teleology and intentionality in a human way. It is to subscribe to the notion that emotional response to music is fundamentally no different from emotional response to anything else.

Although this chapter has been mainly concerned with highlighting the potential of the Sound-Utterance-Context-Narrative model as a framework for future research—particularly if more empirical studies on narrative are undertaken—the main message of the thesis lies not so much within the model itself but within the approach taken to formulate and explore it: that music has a great power to evoke emotions is undeniable, but the source of this power is not embedded deeply within the intricacies and origins of a musical culture;

instead, it derives from listeners' ability to make coherent sense out their environment and their inclination toward assessing this environment continuously with respect to their own well-being. If music is in some way "special", it is perhaps not so much because a listening experience invokes some special cognitive powers, but because musical stimuli are by nature incoherent and incomplete; music is not a self-contained world. A listener must invoke a plethora of analogies, metaphors and memories in order to make coherent sense out of the auditory environment, and the very struggle to achieve coherence can even be an emotional experience in itself. The implications of all this are abundantly clear: if music psychology is to make a leap from sophisticated observation of phenomenology to a coherent understanding of emotional response to music, it must embrace the rich, domain-independent multimedia context in which music is heard and experienced. A quest for an understanding of emotional response to music, then, must be a quest for an understanding of the ways in which listeners attempt to make sense of the world. To repeat once more the central mantra of this thesis, such an understanding can come only from an approach that considers emotional response to music as emotional response to a perfectly ordinary human engagement.

Appendices

Appendix A

Software and Equipment

All three experiments presented in Part Two of the thesis relied on custom-written software for their execution; the second and third experiments also used a custom-built pressure sensor. This appendix, which is primarily intended to help anyone wishing to replicate any of the experiments, provides complete code listings of the software, a circuit diagram for the pressure sensor, and a brief explanation of how each works. Every code listing is preceded by an overview describing its purpose and design; more detailed explanation of its workings and rationale for any idiosyncratic design features are embedded within the listings themselves (in the form of "comments"). Whereas the overviews do not assume any specific programming knowledge, the comments embedded within listings assume familiarity with the relevant programming language (Java, C or C++). All the code listed here is also included on *CD One - Software*, attached to this thesis.

Software used in Chapter Seven

This experiment used *The MuPsych Lab*, experiment design and implementation software written by the author of this thesis (Lavy, 1999). The software is currently undergoing major revisions and will soon be available in its new form from *http://www.mupsych.org*. In the incarnation used for the experiment, *The MuPsych Lab* consisted of two discrete components: a server for controlling experiment sessions (displaying stimuli and collecting results), and a programmers' library. The run-time interface for a *MuPsych Lab* experiment is a Java Applet sitting in a web page; the server software presents trials to participants via the Applet and collects responses in a central repository. Although the software allows many participants to complete trials simultaneously (all that is required is an Internet-connected computer for each partipant), this facility was not used for the experiment presented in Chapter Seven. An experiment that runs in *The MuPsych Lab* environment consists of a Java "Class" that uses the MuPsych Programming Library. This library serves two functions. First, it provides access to the run-time environment, presenting programmers with the tools required to write a MuPsych experiment. Second, it provides a set of pre-built modules for use inside an experiment trial: there is, for example, a module for playing sound files, one for presenting and collecting data from rating scales, and one for forced-choice paradigms. Documentation listing the various features of this Programming Library and explaining how to use them from within Java may be found on *CD One - Software*.

Technically, a programmer wishing to write an experiment for use with MuPsych has to implement an Experiment Class, which is responsible for creating a Trial List, specifying what participant-specific information should be collected before an experiment session starts, and writing the data to a suitable storage medium; a set of functions within the Programmers' Library helps with these tasks. The Experiment Class for the experiment presented in Chapter Seven is presented below; an understanding of it requires familiarity with Java; basic knowledge of the MuPsych Programmers' Library is not a pre-requisite but will help. Anyone intending to use this code is encouraged to contact the author of this thesis for detailed information about how to set up a MuPsych experiment.

```
package org.mupsych.lab.core;
import org.mupsych.lab.server.*;
import org.mupsych.lab.presenter.*;
import java.util.*;
import java.rmi.RemoteException;
import java.io.*; //temporary

public class Experiment
implements SessionServer
{

    private char GROUP;

    /* Constructor - this experiment ordered trials differently
       depending  on whether a participant was allocated to group A
       or group B. The group information is passed to Experiment
       objects on construction in the form of an experiment name.
```

```
      Hence, we extract that data here and set a var
   */
   public Experiment(String expname)
   {
       pExpName = expname;
       if(expname.equals("exp1_groupA"))
       {
           GROUP = 'A';
       }
       else {GROUP = 'B';}
   }

   /* This method constructs a Trial List. Most of the code is
      concerned with generating randomised orders etc. It uses four
      of the MuPsych modules: RatingScale, OptionScale,
      TriggerButton and Sound. Their function should be
      self-explanatory
   */
   public synchronized List getTrialList()
   {

      List trials = new ArrayList();
      try{

       //First create the 3 groups of stimuli and 4 groups of texts
       //Randomise each
       List textsA = new ArrayList();
       textsA.add("Text A");
       textsA.add("Text B");
       textsA.add("Text E");
       Collections.shuffle(textsA);

       List textsB = new ArrayList();
       textsB.add("Text C");
       textsB.add("Text D");
       textsB.add("Text F");
       Collections.shuffle(textsB);

       List textsC = new ArrayList();
       textsC.add("Text G");
       textsC.add("Text J");
       textsC.add("Text K");
       Collections.shuffle(textsC);

       List textsD = new ArrayList();
       textsD.add("Text H");
       textsD.add("Text I");
       textsD.add("Text L");
       Collections.shuffle(textsD);
```

```
List musicA = new ArrayList();
musicA.add("1.WAV");
musicA.add("6.WAV");
musicA.add("9.WAV");
Collections.shuffle(musicA);

List musicB = new ArrayList();
musicB.add("2.WAV");
musicB.add("3.WAV");
musicB.add("4.WAV");
Collections.shuffle(musicB);

List musicC = new ArrayList();
musicC.add("7.WAV");
musicC.add("10.WAV");
musicC.add("12.WAV");
Collections.shuffle(musicC);

List[] textGroups = { textsA, textsB, textsC, textsD };
List[] musicGroups = { musicA, musicB, musicC };


//create the Trials. Each text type must go with each music
//type each time we grab text or music from a group, we
//take the first one in the list that has not been used
//before. When we run out of music, we start again from the
//top of the list

//for each textGroup
for (int i=0; i < textGroups.length; i++)
{
    //for each musicGroup
    for(int j=0; j < musicGroups.length; j++)
    {
        //find which text and music to use
        String textToUse =
           (String)textGroups[i].get(j%textGroups.length);
        String musicToUse =
           (String)musicGroups[j].get(i%musicGroups.length);

        //List for compiling presentables
        List pres = new ArrayList();

        //Create Presenters and get refs to relevant
        //Control Points and Announcements
        TitleDisplay st = new TitleDisplay("StartTrial");
        st.setText(textToUse);
        ControlPoint stShow =
            (ControlPoint)st.getControlPoints().first();
        ControlPoint stHide =
```

```
        (ControlPoint)st.getControlPoints().last();
    pres.add(st);

    TriggerButton pm = new TriggerButton("PlayExtract");
    pm.setButtonText("Start Trial");
    ControlPoint pmShow =
        (ControlPoint)pm.getControlPoints().last();
    Announcement pmClicked =
        (Announcement)pm.getAnnouncements().first();
    pres.add(pm);

    Sound ex = new Sound("Extract");
    ex.setMedia(musicToUse);
    ControlPoint exPlay =
        (ControlPoint)ex.getControlPoints().first();
    Announcement exStopped =
        (Announcement)ex.getAnnouncements().last();
    pres.add(ex);

    TriggerButton fr = new TriggerButton("StopReading");
    fr.setButtonText("I have finished reading");
    ControlPoint frShow =
        (ControlPoint)fr.getControlPoints().last();
    Announcement frClicked =
        (Announcement)fr.getAnnouncements().first();
    pres.add(fr);

    RatingScale qu1 = new RatingScale("QuTextMov");
    qu1.setLabel("How moving did you find this story?");
    qu1.setLeftLabel(
       "As unmoving as any that I have ever read");
    qu1.setRightLabel(
       "As moving as any that I have ever read");
    qu1.setNumPoints(7);
    ControlPoint qu1Enable =
        (ControlPoint)qu1.getControlPoints().last();
    Announcement qu1Done =
        (Announcement)qu1.getAnnouncements().first();
    pres.add(qu1);

    OptionScale qu2 = new OptionScale("QuTextVal");
    qu2.setLabel(
     "If you found the story moving, was it positively"+
     " moving, negatively moving or both/neither "+
     "(positive = e.g. Made me happy, amused, etc...; "+
     "negative = e.g. made me sad, morose, contemplative"+
     "...; neither = I was not moved by the story; "+
     "don't know)");
    qu2.setNoDataValue(-99);
    OptionScale.Option[] vals = {
```

```
    new OptionScale.Option("Positive", 1),
    new OptionScale.Option("Negative", -1),
    new OptionScale.Option("Neither/Both", 0),
    new OptionScale.Option("Don't know", 0)
};
qu2.setOptions(vals);
ControlPoint qu2Enable =
    (ControlPoint)qu2.getControlPoints().last();
Announcement qu2Done =
    (Announcement)qu2.getAnnouncements().first();
pres.add(qu2);

OptionScale qu3 = new OptionScale("QuReadBefore");
qu3.setLabel("Have you read this story before?");
qu3.setNoDataValue(-99);
OptionScale.Option[] vals2 =
    { new OptionScale.Option("Yes", 1),
      new OptionScale.Option("No", 0)
    };
qu3.setOptions(vals2);
ControlPoint qu3Enable =
    (ControlPoint)qu3.getControlPoints().last();
Announcement qu3Done =
    (Announcement)qu3.getAnnouncements().first();
pres.add(qu3);

RatingScale qu4 = new RatingScale("QuMusMov");
qu4.setLabel("How moving did you find this music?");
qu4.setLeftLabel(
    "As unmoving as any that I have ever heard");
qu4.setRightLabel(
    "As moving as any that I have ever heard");
qu4.setNumPoints(7);
ControlPoint qu4Enable =
    (ControlPoint)qu4.getControlPoints().last();
Announcement qu4Done =
    (Announcement)qu4.getAnnouncements().first();
pres.add(qu4);

OptionScale qu5 = new OptionScale("QuMusVal");
qu5.setLabel(
 "If you found the music moving, was it positively" +
 " moving, negatively moving or both/neither "+
 "(positive = e.g. Made me happy, amused, etc...;"+
 " negative = e.g. made me sad, morose, contemplative"+
 "...; neither = I was not moved by the music; "+
 "don't know)");
qu5.setNoDataValue(-99);
qu5.setOptions(vals);
ControlPoint qu5Enable =
```

```
    (ControlPoint)qu5.getControlPoints().last();
Announcement qu5Done =
    (Announcement)qu5.getAnnouncements().first();
pres.add(qu5);

OptionScale qu6 = new OptionScale("QuHeardBefore");
qu6.setLabel(
   "Have you heard this piece of music before?");
qu6.setNoDataValue(-99);
qu6.setOptions(vals2);
ControlPoint qu6Enable =
    (ControlPoint)qu6.getControlPoints().last();
Announcement qu6Done =
    (Announcement)qu6.getAnnouncements().first();
pres.add(qu6);

OptionScale qu7 = new OptionScale("QuSeparate");
qu7.setLabel(
 "Did you feel that reading the story and listening "+
 "to the music were very separate experiences (i.e. "+
 "your emotional response to the text was separate "+
 "from your response to the music), or did you "+
 "respond to music and text as an integrated whole?"
);
OptionScale.Option[] vals3 = {
     new OptionScale.Option("Completely separate", 4),
     new OptionScale.Option("Slightly separate", 3),
     new OptionScale.Option("Fairly integrated", 2),
     new OptionScale.Option("Completely integrated", 1)
};
qu7.setNoDataValue(-99);
qu7.setOptions(vals3);
ControlPoint qu7Enable =
     (ControlPoint)qu7.getControlPoints().last();
Announcement qu7Done =
     (Announcement)qu7.getAnnouncements().first();
pres.add(qu7);

TriggerButton nt = new TriggerButton("NextTrial");
nt.setButtonText("Click here for next trial");
ControlPoint ntEnable =
      (ControlPoint)nt.getControlPoints().last();
Announcement ntClicked =
      (Announcement)nt.getAnnouncements().first();
pres.add(nt);


//MuPsych trials consist of a set of Presenters,
//all of which are linked together by preconditions;
//a Presenter is displayed to a participant when all
```

```
//of its preconditions have been met.
//These are defined below
Map pre = new HashMap();

//TrialStart -> stShow and pmShow
Set s = new HashSet();
s.add(stShow);
s.add(pmShow);
pre.put(new TrialStartAnnouncement(), s);

//pmClicked -> exPlay and frShow
s = new HashSet();
s.add(exPlay);
s.add(frShow);
pre.put(pmClicked, s);

if(GROUP == 'A')
{
    //frClicked -> qu1Enable
    s = new HashSet();
    s.add(qu1Enable);
    pre.put(frClicked, s);

    //exStopped -> qu1Enable
    s = new HashSet();
    s.add(qu1Enable);
    pre.put(exStopped, s);

    //qu1Done -> qu2Enable
    s = new HashSet();
    s.add(qu2Enable);
    pre.put(qu1Done, s);

    //qu2Done -> qu3Enable
    s = new HashSet();
    s.add(qu3Enable);
    pre.put(qu2Done, s);

    //qu3Done -> Qu4Enable
    s = new HashSet();
    s.add(qu4Enable);
    pre.put(qu3Done, s);

    //qu4Done -> Qu5Enable
    s = new HashSet();
    s.add(qu5Enable);
    pre.put(qu4Done, s);

    //qu5Done -> Qu6Enable
    s = new HashSet();
```

```
        s.add(qu6Enable);
        pre.put(qu5Done, s);

        //qu6Done -> Qu7Enable
        s = new HashSet();
        s.add(qu7Enable);
        pre.put(qu6Done, s);

    }
    else
    {
        //frClicked -> qu4Enable
        s = new HashSet();
        s.add(qu4Enable);
        pre.put(frClicked, s);

        //exStopped -> qu4Enable
        s = new HashSet();
        s.add(qu4Enable);
        pre.put(exStopped, s);

        //qu4Done -> qu5Enable
        s = new HashSet();
        s.add(qu5Enable);
        pre.put(qu4Done, s);

        //qu5Done -> qu6Enable
        s = new HashSet();
        s.add(qu6Enable);
        pre.put(qu5Done, s);

        //qu6Done -> Qu1Enable
        s = new HashSet();
        s.add(qu1Enable);
        pre.put(qu6Done, s);

        //qu1Done -> Qu2Enable
        s = new HashSet();
        s.add(qu2Enable);
        pre.put(qu1Done, s);

        //qu2Done -> Qu3Enable
        s = new HashSet();
        s.add(qu3Enable);
        pre.put(qu2Done, s);

        //qu3Done -> Qu7Enable
        s = new HashSet();
        s.add(qu7Enable);
        pre.put(qu3Done, s);
```

```
            }

            //qu7Done -> ntEnable
            s = new HashSet();
            s.add(ntEnable);
            pre.put(qu7Done, s);

            //ntClicked -> EndOfTrial
            s = new HashSet();
            s.add(new TrialEndControlPoint());
            pre.put(ntClicked, s);

            trials.add(new Trial(pre, pres,new Response(
             new TrialID(textToUse + ":" + musicToUse),0)));

        }

    }

    //randomise Trial List!
    Collections.shuffle(trials);
    Iterator it = trials.iterator();
    for (int i = 0; it.hasNext(); i++)
    {
        ((Trial)it.next()).getResponse().setPlayOrder(i + 1);
    }


    }catch(Exception e){e.printStackTrace();}

    finally{return Collections.unmodifiableList(trials);}
}

/* This method creates an experiment session for a partcipant.
   This involves instantiating a SessionImpl.
*/
public synchronized Session createSession(String clientIP)
throws RemoteException
{
    Random rnd = new Random();
    String ID = String.valueOf(rnd.nextInt());
    Session sn = new SessionImpl(this,getName(),clientIP,ID);
    return sn;
}

/* This method returns a PIG (personal information gatherer), used
   to collect participant-specific data that is not part of a
   trial
*/
```

```
    public synchronized Pig getPig()
    {
        String str;
        String instr = "";
   try
    {
       BufferedReader br = new BufferedReader(
            new FileReader
               ("/tmp/instr" + String.valueOf(GROUP) + ".txt"));
      while ((str = br.readLine()) != null)
       {
               instr += str;
               instr += "\n\n";
            }
        }
        catch(Exception e){
            instr = "Ooops. Couldn't find instructions!";
         e.printStackTrace();
        }

        //build the pig
        List items = new ArrayList();
        items.add(new PigItem(
           "Group Number: - please ask experimenter",
                String.valueOf(GROUP)));
            items.add(new PigItem("What is your age?",""));
            items.add(new PigItem("What is your gender (M/F)?",""));
            items.add(new PigItem(
                 "Do you regularly listen to music?",""));
            items.add(new PigItem(
                "How many hours (approximately) do you spend "+
                "listening to music per week?",""));
            items.add(new PigItem("Do you read for pleasure?",""));
            items.add(new PigItem(
                "How many hours (approximately) do you spend "+
                "reading for pleasure per week?",""));
            return new Pig(instr,items);
    }


    /* A debriefing message displayed to participants once the
       experiment session is over
    */
    public String getDebrief()
    {
        return  "\n\nThe experiment is now over. "+
          "Thank you very much for your participation.\n\n"+
          "You may now take off your headphones. please inform the "+
          "experimenter that you have finished.";
    }
```

```
    /* MuPsych allows Experiments to check (using arbitrary criteria)
       whether answers to PIG questions should be accepted as valid.
       We are not using this facility here; hence return null.
    */
    public PigError validatePig(Pig thePig)
    {
        return null;
    }

    /*  Returns the registered name of the Experiment */
    public String getName()
    {
        return pExpName;
    }

    /* This method is passed the Session Data once a session
       is complete. It is the responsibility of this method to
       write the data out to file - or in this case STDOUT.
    */
    public synchronized void depositSessionData(SessionData sd)
    {
        System.out.println(sd.toString());
    }

    private String pExpName;
    private boolean piggy;

}
```

In order to run the code, it should be compiled with any Java compiler and then added to the Experiment Loader. For further details, please contact the author of this thesis.

Software used in Chapter Eight

The experiment presented in Chapter Eight required software that could provide accurate timing of button clicks and could interact with a custom-built hardware device, namely the pressure sensor. As Java is a highly unsuitable platform for accomplishing either of these tasks, *The MuPsych Lab* could not be used here. Hence, I wrote a self-contained program to run the experiment, using a combination of the C and C++ programming languages. The program runs on the Linux platform; it uses the QT toolkit (*http://www.trolltech.com*) for windowing and multi-threading, and the Parapin library (*http://www.circlemud.org/~jelson/software/parapin*) to control the PC parallel port, which is needed for interacting with the pressure sensor. Both

QT and Parapin are available at no cost from their respective websites. The software written for this experiment consists of six C/C++ source files and their associated headers. These are listed in full below.

The Parapin library reads from and write to the parallel port directly; under Linux, these operations require a program to have special privileges. Hence, the first thing the program must do is enable these privileges and initialise the parallel port; this is accomplished in MAIN.C:

```
/* Experiment Two - Stimulus Presenter and Response gatherer
 * ------------------------------------------------------------
 * main.c:
 *     This code (which must run as root) inits the parallel port
 *     for the parapin library. Having done this, we drop root
 *     privileges and call qtmain.
 */

#include <parapin.h>
#include <unistd.h>
#include <stdio.h>
#include <sys/io.h>

int qtmain(int argc, char **argv);

int main (int argc, char **argv)
{

#ifndef NO_PSENSOR
    //give this process and any fork()ed or exec()ed processes
    // full permissions for all I/O.
    if (iopl(3) != 0){
        printf("Could not set I/O permissions\n");
        exit(1);
    }
    //init parapin
    if (pin_init_user(LPT1) != 0){
        printf("Could not initialise parallel port.\n");
        exit(2);
    }


#endif

    //drop root privileges
    if (setuid(getuid()) != 0){
        printf("Could not set effective userID.\n");
        exit(3);
```

```
    }

//pass control to QT
    return qtmain(argc, argv);
}
```

If initialisation is successful, control is passed to a function in QTMAIN.CPP; here, QT is initialised and an Experiment object created.

```
/* Experiment Two - Stimulus Presenter and Response Gatherer
 * qtmain.cpp - initialises the QT Environment */

#include <qapplication.h>
#include <unistd.h>
#include <stdio.h>
#include "Experiment.h"

extern "C" int qtmain (int argc, char **argv)
{
    QApplication a(argc, argv);
    if (argc != 2){ //only the interesting args should remain now
        printf("I require an output file name.\n");
        exit(4);
    }
    Experiment exp(argv[1]);
    return a.exec();
}
```

The Experiment class contains the code that actually runs the experiment. It is responsible for ordering trials and, with the help of the ExpDisplay class, displaying stimuli and gathering responses. The definition of Experiment and its implementation (EXPERIMENT.H and EXPERIMENT.CPP respectively) are printed below:

```
/* Experiment - class that encapsulates the experiment.
 * controles sesssion flow, stimuli, responses, etc */

#ifndef _Experiment_
#define _Experiment_

#include <qobject.h>
#include <qtimer.h>
#include <qstring.h>
#include <fstream.h>
#include <iostream.h>

#include "ExpDisplay.h"
#include "PressureSensor.h"
```

```
class Experiment : public QObject,QThread
{
    Q_OBJECT

    public:
        Experiment(char* outputfile);
    public slots:
        void hearButtonClick(int btn); //btn identifies the button
        void updateDialValue();
    private:
        ExpDisplay* theDisplay;
        PressureSensor* theSensor;
        QTimer* theTimer;
        QTime* timeObject;
        void doNextThing(int btn);
        char** instructions;
        char** propositions; //for proposition[]
        char** musics; //for musics[]
        char** questions; //for questions[]
        int currentSection; //keep tabs for play time
        int currentTrial; //ditto
        ofstream resultsStream;
        char* fileToString(char* filename);
        int fileToStringArray(char** &array, char* filename);
        int doNextProposition();
        int doNextQuestion(int btn);
        void run(); //QThread override
        void randomiseNullDelimCharArray(char** array);
};

#endif /* _Experiment_ */

/* Implementation for Experiment class.
 * Experiment flow:
 *  (1) Display instructions
 *  (2) Do Stimulus Trial block
 *  (3) Display instructions about the non-computer filler task
 *  (4) Display instructions about the final part
 *  (5) Do Question Trial block
 *  (6) Debrief
 *
 * Each stimulus Trial consists of:
 *  (1) Display proposition with 'Click when you are ready to start'
 *      button; start sensor polling
 *  (2) Continue to display proposition with DISABLED 'Click for next
 *      trial' button; play music; record sensor
 *  (3) Write out proposition, music & sensor data
 *  (4) Enable 'Click for next trial' button
 *
 * The order for music and proposition are independently permuted.
```

```
 * Each question stimulus consists of:
 *  (1) Display question with 'true' and 'false' buttons; start timer
 *  (2) Write out which button was pressed and timing data
 */

#include "Experiment.h"
#include <qapplication.h>
#include <sys/stat.h>
#include <unistd.h>
#include <iomanip.h>
#include <qdatetime.h>
#include <string.h>
#include <stdlib.h>
#include <time.h>

/* defines for filenames */
#define IFILE_1 "../../texts/stimulus_instr.txt"
#define IFILE_2 "../../texts/filler_instr.txt"
#define IFILE_3 "../../texts/question_instr.txt"
#define IFILE_4 "../../texts/debrief_instr.txt"
#define QUESFILE "../../texts/facts.txt"  //list of questions
#define MUSIFILE "../../texts/musics.txt" //list of music FILENAMES
#define PROPFILE "../../texts/propositions.txt" //list of propositions

#define SOUND_PLAYER "/usr/bin/mpg123 " //NB the space!

Experiment::Experiment(char* outputfile)
{
    currentSection = currentTrial = 0;
    theSensor = new PressureSensor();
    theDisplay = new ExpDisplay(NULL, "Display");
    theTimer = new QTimer(theDisplay);
    timeObject = new QTime();
    connect(theTimer, SIGNAL(timeout()),
             this, SLOT(updateDialValue()));
    connect(theDisplay, SIGNAL(buttonClicked(int)),
             this, SLOT(hearButtonClick(int)));

    //import instructions or quit with fatal error
    instructions = new char*[4];
    for (int i=0; i<4;i++) instructions[i] = NULL;
    if (    (instructions[0] = fileToString(IFILE_1)) == NULL ||
            (instructions[1] = fileToString(IFILE_2)) == NULL ||
            (instructions[2] = fileToString(IFILE_3)) == NULL ||
            (instructions[3] = fileToString(IFILE_4)) == NULL
        ){
            qFatal("Problem: one of the instruction files is broken.\n");
    }

    //read in propositions and music files;
```

```
    //make sure they are of equal number
    if (fileToStringArray(propositions, PROPFILE)
         != fileToStringArray(musics, MUSIFILE)){
        qFatal("Problem: proposition or music file not usable.");
    }

    //did it work?
    if (propositions == NULL){
        qFatal("Problem: propositions == NULL.");
    }
    if (musics == NULL){
        qFatal("Problem: musics == NULL");
    }

    //read in question file
    fileToStringArray(questions, QUESFILE);
    if (questions == NULL){
        qFatal("Problem: questions == NULL");
    }

    //check music files are all statable
    struct stat fi;
    for (int i=0; musics[i] != NULL; i++){
        if (stat(musics[i],&fi)){
            qFatal("Problem: Music file %s does not exist", musics[i]);
        }
    }

    //set up random seed
    srand((unsigned int)time(NULL));

    //randomise order of musics, questions and propositions
    randomiseNullDelimCharArray(musics);
    randomiseNullDelimCharArray(questions);
    randomiseNullDelimCharArray(propositions);

    //open a resultsfile for writing
    resultsStream.open(outputfile, ios::out|ios::noreplace|ios::stdio);
    if (resultsStream.bad()){
        qFatal("Problem: Could not open %s for writing.\n", outputfile);
    }
    resultsStream << "Experiment Session started on "
                  << QDateTime::currentDateTime().toString().latin1()
                  << "\n";

    //initialise environment
    theDisplay->show();
    qApp->setMainWidget(theDisplay);
    doNextThing(0);
```

```
}

/* SLOT */
void Experiment::hearButtonClick(int btn)
{
    doNextThing(btn);
}

/* Here, events are heard and we call the relevant routine
 * depending on where we are in the experiment.
 */
void Experiment::doNextThing(int btn)
{
    switch (currentSection){ //what part of the experiment are we on?
        case 0: //initial instructions
            theDisplay->setInstructionMode(
             instructions[0],
             "Click here to start experiment");

            currentSection++;
        break;

        case 1: //propositions and sensor
            if (doNextProposition() == 0){
                currentSection++;
                doNextThing(0);
            }
        break;

        case 2: //filler task instructions
            theDisplay->setInstructionMode(
              instructions[1],
              "Do not click until you have completed this task");
            currentSection++;
        break;

        case 3: //question part instructions
            theDisplay->setInstructionMode(
              instructions[2],
              "Click here to start the question sequence");
            currentSection++;
        break;

        case 4: //question trials
            if (doNextQuestion(btn) == 0){
                currentSection++;
                doNextThing(0);
            }
        break;
```

```
        case 5: //debrief instructions
            theDisplay->setInstructionMode(
              instructions[3],
              "Click here to finish.");
            currentSection++;
        break;

        default: //experiment has ended
            resultsStream << "__END OF SESSION__";
            resultsStream.flush();
            resultsStream.close();
            qApp->exit(0);
    }
}


/* Runs the next question, based on currentTrial. Returns -1 if
 * there is something to do. Once there are no more questions to do,
 * currentTrial is reset to zero and zero is returned.
 */
int Experiment::doNextProposition()
{
    if (currentTrial == 0){ //we're on first proposition
        //init sensor reading
        theSensor->startReading();
        theTimer->start(150, FALSE); //150msec
    }

    //Now for the actual trial. We need to pop up proposition
    //with a click here to continue & then actually start
    //recording / music. Therefore, every SECOND time this is called
    //is a new trial. currentTrial is even means pop up new proposition;
    // odd means play music. As propositions[] and musics[] are not
    // step 2, we divide currentTrial by 2 to get array index

    if (! (currentTrial & 1)){ //it's even
            int index = currentTrial >> 1;
        if (propositions[index] == NULL){ //NO MORE TRIALS
            theTimer->stop();
            theSensor->stopReading();
            currentTrial = 0; //reset it for questions
            return 0;
        } else { //more trials
            resultsStream << "\nProposition: "
                          << propositions[index] << "\n";
            theDisplay->setStimulusMode(
              propositions[index],
              "Click here to start the music");
        }
        currentTrial++;
```

```
    } else {
        //it's odd, so do recording and playing music - this happens
        // in run() because we want it to happen in a background thread
        theDisplay->disableButtons();
        start();
    }
    return 1;
}

/* The play music thread. This is called by doNextProposition() when
 * it is time for music to be played. It is not inherently thread-safe,
 * but this is handled by the caller, which invokes
 * theDisplpay->disableButtons() ensureing that nothing else can
 * happen while we're going
 */
void Experiment::run()
{
        //it's odd
        theDisplay->disableButtons();
        int index = ((currentTrial - 1) >> 1);

        //prepare commandline for media player
        char* cmd = new char[sizeof(SOUND_PLAYER) +
                             strlen(musics[index]) +
                             sizeof(" 2>/dev/null")];
        strcpy(cmd, SOUND_PLAYER);
        strcat(cmd, musics[index]);
        strcat(cmd, " 2>/dev/null");

        //record && play
        theSensor->startRecording();
        system(cmd);
        QThread::msleep(500);
        theSensor->stopRecording();
        delete[] cmd;
        cmd = NULL;

        //writeout data
        resultsStream << "Music: " << musics[index] << "\n";
        resultsStream << "Sensor Data: \n";
        theSensor->writeBuffer(resultsStream);
        resultsStream << "\n";

        currentTrial++;
        theDisplay->setStimulusMode(
          propositions[index],
          "Click here to continue");
        theDisplay->enableButtons();
}
```

```
/* Runs the next question, based on currentTrial. Returns -1 if
 * there is something to do. Once there are no more questions to do,
 * currentTrial is reset to zero and zero is returned
 */
int Experiment::doNextQuestion(int btn)
{

    if (currentTrial == 0){ //first time we're called
        resultsStream << "\nQuestions: <qu>\t<msec>\t<button>\n";

    } else { //not the first time we're called so save previous
        resultsStream << timeObject->elapsed();
        resultsStream << "\t" << (btn ? "True" : "False") << "\n";
    }

    if (questions[currentTrial] == NULL){ //no more questions
        currentTrial = 0;
        return 0;

    } else { //there are...
        resultsStream << questions[currentTrial] << "\t";
        theDisplay->disableButtons();
        theDisplay->setQuestionMode(
          questions[currentTrial],
          "True", "False");
        currentTrial++;

        theDisplay->enableButtons();
        timeObject->start(); //reset timer
        return -1;
    }
}


/* SLOT */
void Experiment::updateDialValue()
{
    theDisplay->updateDial(theSensor->getCurrentValue());
}

/* Grab the contents of a file and put them into a single
 * null-terminated string; returns pointer to the newly
 * created string or null if it fails*/
char* Experiment::fileToString(char* filename)
{
    //allocate requisite memory
    struct stat fi;
    if (stat(filename,&fi) != 0){
        return NULL;
```

```
    }
    char* buffer = new char[fi.st_size + 1];

    //open file
    ifstream ifs(filename, ios::in);
    if (ifs.bad()){
        return NULL;
    }
    ifs.unsetf(ios::skipws); //do not ignore '\n' and ' '

    //fill buffer && return
    for (int i = 0; i < fi.st_size; i++){
        ifs >> buffer[i];
    }
    buffer[fi.st_size] = '\0';
    ifs.close();
    return buffer;
}

/* Grab the contents of a file and put them into a string array,
 * one string per line of the original file (i.e. split on '\n').
 * Memory is allocated as required; final element of array is always
 * NULL so the end can be determined. Returns the number of elements
 * in the array (_INCLUDING_ the final NULL element), or in case of
 * problem, it just returns 0 and sets char** array to NULL.
 */
int Experiment::fileToStringArray(char** &array, char* filename)
{
    //grab to temporary buffer
    char* tmp;
    if ((tmp = fileToString(filename)) == NULL){
        array = NULL;
        return 0;
    }

    //work out no. lines required, and allocate memory for pointers
    int numlines = 0;
    for (int i=0; tmp[i] != '\0'; i++){
        if (tmp[i] == '\n') numlines++;
    }
    array = new char*[numlines + 1];

    //allocate the char arrays (one per line) and populate
    int which = 0; //current line
    char* s = tmp; //pointer to start of it
    for (int i=0; tmp[i] != '\0'; i++){
        if (tmp[i] == '\n'){
            tmp[i] = '\0'; //replace '\n' with '\0'
            array[which] = s; //save pointer to start
            s = &tmp[i+1]; //s points to start of next string
```

```
                which++;
            }
        }
        array[numlines] = NULL;
        return numlines;
}

/* Takes an array of char*s and permutes the order.
 * Assumes that the final element of the array
 * (which will NOT be moved from its position) is NULL)
 */
void Experiment::randomiseNullDelimCharArray(char** array)
{
        //how long is array?
        int arraylen = 0;
        while (array[arraylen] != NULL){
            arraylen++;
        }

        //perform the randomisation, swapping i with rnd(i to arraylen-1)
        char* tmp;
        int rnd;
        for (int i=0; i < arraylen; i++){
            rnd = 1 + (int)(((arraylen-1)*(double)rand())/(RAND_MAX+1.0));
            tmp = array[rnd];
            array[rnd] = array[i];
            array[i] = tmp;
        }
}
```

The ExpData class is primarily responsible for creating and controlling window controls; it presents Experiment with an abstraction of the various QT widgets used to present stimuli and gather responses. Its definition can be found in EXPDISPLAY.H and its implementation in EXPDISPLAY.CPP:

```
/* ExpDisplay.h - The display window for Experiment Two */

#ifndef _ExpDisplay_
#define _ExpDisplay_

#include <qwidget.h>
#include <qpushbutton.h>
#include <qdial.h>
#include <qlabel.h>
#include <qlayout.h>
#include <qtextview.h>

#define BTN_BUTTON_0 0
```

```
#define BTN_BUTTON_1 1

class ExpDisplay : public QWidget
{
    Q_OBJECT
    public:
        ExpDisplay(QWidget* parent, const char* name);
        void setInstructionMode(char* instructions, char* buttontxt);
        void setStimulusMode(char* stimulus, char* buttontxt);
        void setQuestionMode(char* question, char* btxt_0, char* btxt_1);
        void enableButtons();
        void disableButtons();
        void updateDial(int val);
    private slots:
        void button_0_clicked();
        void button_1_clicked();
    signals:
        void buttonClicked(int button);
    private:
        QGridLayout* layoutMgr;
        QTextView* longText;
        QLabel* text;
        QPushButton* button_0;
        QPushButton* button_1;
        QDial* dial;
        int currentMode; //flag the display status
        void hideAllWidgets();
        void initLayout();
};

#endif /* _ExpDisplay_ */
```

```
/* ExpDisplay.cpp - implementation for Display class
 * the display widget for Experiment Two */

#include "ExpDisplay.h"
#include <qfont.h>
#include <qcolor.h>
#include <qpalette.h>
#include <qtimer.h>

#ifndef NULL
    #define NULL 0
#endif

#define M_NO_MODE 0
#define M_INSTRUCTION_MODE 1
#define M_STIMULUS_MODE 2
#define M_QUESTION_MODE 3
#define WND_WIDTH 800
#define WND_HEIGHT 500
```

```
ExpDisplay::ExpDisplay(QWidget *parent, const char *name)
        : QWidget (parent, name)
{
    currentMode = M_NO_MODE;
    setMinimumSize(WND_WIDTH,WND_HEIGHT);
    layoutMgr = NULL;
    text = new QLabel(this);
    text->setPalette(QPalette(QColor(255,255,255)));
    text->setFrameStyle(QFrame::Panel | QFrame::Sunken);
    longText = new QTextView(this);
    longText->setHScrollBarMode(QScrollView::AlwaysOff);

    //set up layout manager
    initLayout();

    //set up widgets
    QFont* buttonFont = new QFont("Helvetica", 18, QFont::Bold);

    button_0 = new QPushButton("Button 0", this, "button_0");
    button_0->setFont(*buttonFont);
    button_0->hide();

    button_1 = new QPushButton("Button 1", this, "button_1");
    button_1->setFont(*buttonFont);

    dial = new QDial(this);
    dial->setMinValue(0);
    dial->setMaxValue(127);
    dial->setEnabled(FALSE);

    connect(button_0, SIGNAL(clicked()),
            this, SLOT(button_0_clicked()));
    connect(button_1, SIGNAL(clicked()),
            this, SLOT(button_1_clicked()));
}


/* Prepare the window for displaying instructions (if not done
 * already and display the instructions in the window with a
 * disabled dismiss button
 */
void ExpDisplay::setInstructionMode(char* instructions, char* buttontxt)
{
    hideAllWidgets();
    if (currentMode != M_INSTRUCTION_MODE){ //need to set up window
        initLayout();
        layoutMgr->addMultiCellWidget(longText,0,0,0,3);
        layoutMgr->addMultiCellWidget(button_0,1,1,0,3);
    }
```

```
    longText->setText(instructions);
    button_0->setText(buttontxt);

    longText->show();
    button_0->show();
}

/* Prepare the window for displaying the stimuli (if not done already)
 * and display the passed stimulus text in the window with a disabled
 * 'next' button
 */
void ExpDisplay::setStimulusMode(char* stimulus, char* buttontxt)
{
    hideAllWidgets();
    if (currentMode != M_STIMULUS_MODE){ //need to set up window
        initLayout();
        layoutMgr->addMultiCellWidget(text,0,0,0,3);
        layoutMgr->addMultiCellWidget(button_0,1,1,0,2);
        layoutMgr->addWidget(dial,1,3);
        text->setAlignment(AlignHCenter| AlignVCenter | WordBreak);
        text->setFont(QFont("Helvetica", 24, QFont::Normal));
    }
    text->setText(stimulus);
    button_0->setText(buttontxt);

    text->show();
    button_0->show();
    button_0->repaint();
    dial->show();
}

/* Prepare the window for asking questions (if not done already)
 * and display the passed question text in the window with disabled
 * 'True' and 'False' buttons
 */
void ExpDisplay::setQuestionMode(char* question,
                                 char* btxt_0,
                                 char* btxt_1)
{
    hideAllWidgets();
    if (currentMode != M_QUESTION_MODE){ //need to set up window
        initLayout();
        layoutMgr->addMultiCellWidget(text,0,0,0,3);
        layoutMgr->addMultiCellWidget(button_0,1,1,0,1);
        layoutMgr->addMultiCellWidget(button_1,1,1,2,3);
        layoutMgr->addWidget(button_1,1,2);
        layoutMgr->addWidget(button_0,1,1);
        text->setAlignment(AlignHCenter| AlignVCenter | WordBreak);
        text->setFont(QFont("Helvetica", 24, QFont::Normal));
    }
```

```
    text->setText(question);
    button_0->setText(btxt_0);
    button_1->setText(btxt_1);

    text->show();
    button_0->show();
    button_1->show();
}

/* make buttons clickable */
void ExpDisplay::enableButtons()
{
    button_0->setEnabled(TRUE);
    button_1->setEnabled(TRUE);
}

/* make buttons non-clickable */
void ExpDisplay::disableButtons()
{
    button_0->setEnabled(FALSE);
    button_1->setEnabled(FALSE);
}

/* set dial value */
void ExpDisplay::updateDial(int val)
{
    dial->setValue(val);
}


/* Hide all widgets */
void ExpDisplay::hideAllWidgets()
{
        text->hide();
        longText->hide();
        button_0->hide();
        button_1->hide();
        dial->hide();
}

/* Return a new layoutmanager (clean slate) */
void ExpDisplay::initLayout()
{
    if (layoutMgr != NULL){
        delete layoutMgr;
        layoutMgr = NULL;
    }

    layoutMgr = new QGridLayout(this, 2, 4);
    layoutMgr->setMargin(WND_WIDTH >> 5); //margin=1/32nd total
```

```
    layoutMgr->setRowStretch(0,4); // top row:bottom row = 4:1
    layoutMgr->setRowStretch(1,1); //
}

/* private slot implementation - pass on data with params */
void ExpDisplay::button_0_clicked()
        {emit buttonClicked(BTN_BUTTON_0);}
void ExpDisplay::button_1_clicked()
        {emit buttonClicked(BTN_BUTTON_1);}
```

Data from the pressure sensor has to be collected asynchronously (i.e. in a different thread of execution from the rest of the program). In addition, memory for storing the data as a trial progresses has to be allocated and de-allocated dynamically, and converted to a format suitable for writing to a results file. These operations are encapsulated within the PressureSensor class, defined in PRESSURESENSOR.H and implemented in PRESSURESENSOR.CPP:

```
/* PressureScale.h
 * A QObject whose role in life is to provide data from the
 * PressureSensor It emits a stream of signals to represent the
 * position of the PressureSensor at any given time.
 * Once asked to start emitting (with startEmit()), a constant
 * stream of signals will be emitted - one every 50ms - until
 * stopEmit() is called. The value emitted with the signal corresponds
 * to the value of the pressure scale. A call to startRecord()
 * (which is independent of startEmit()) will start saving
 * the data to an int[]. Memory is allocated dynamically,
 * but a hard-coded limit stops it getting _TOO_ ridiculous!
 */

#ifndef _PressureScale_
#define _PressureScale_

#include <qobject.h>
#include <qthread.h>
#include <iostream.h>

#define BLOCK_SZ 150

class PressureSensor : public QObject, QThread
{
    Q_OBJECT

    public:
        PressureSensor();
        ~PressureSensor();
        int getCurrentValue(); //returns currently buffered sample
        int writeBuffer(ostream &os); //writes current buffer to osl
    public slots:
```

```
        void startReading();  //start sampling sensor
        void stopReading();   //stop sampling sensor
        void startRecording();//record sensor status to buffer
        void stopRecording(); //stop recording sensor status to buffer
    protected:
        void run();  //background thread
    private:
        PressureSensor(const PressureSensor&); //no usable copy constr.
        int isRecording;  //flag to toggle buffering
        int shouldStopRecording; //flag to tell poller to stop buffering
        int isReading;  //flag to toggle current poll state
        int currentValue;  //current sample value (or -1);
        struct bufBlock { //block of datapoints with pointer to next
            char dpoint[BLOCK_SZ];
            bufBlock* next;
        } *firstBufBlock;//the initial bufBlock
        void deleteBufferBlock(bufBlock* bb);//de-allocate mem
        void writeBuffer(ostream& os, bufBlock* bb);//writes out bufBlocks
        QWaitCondition waitcond; //for safe stopRecording
};

#endif /* _pressureScale_ */

/* PressureScale.cpp - Implementation of the PressureScale class
 * defined in PressureScale.h
 *
 * NOTE: THE USE OF PARAPIN _REQUIRES_ THAT THE LIBRARY IS INITIALISED
 * BY THE ROOT USER. This is done in main.cpp, after which
 * root privileges are immediately dropped.
 */

#include "PressureSensor.h"
#include "sensorguts.h"
#include <qapplication.h>


#define EMISSION_WAIT 100

/* default constructor */
PressureSensor::PressureSensor()
{
    isRecording = isReading = shouldStopRecording = 0;
    currentValue = -1;
    firstBufBlock = NULL;
}

PressureSensor::~PressureSensor()
{
    //free buffer memory
    if (firstBufBlock != NULL){
        deleteBufferBlock(firstBufBlock);
```

```
    }
}

/* Switch on sensor sampling */
void PressureSensor::startReading()
{
    if (! isReading){
        isReading = 1;
        QThread::start();
    }
}

/* Stop sampling sensor*/
void PressureSensor::stopReading()
{
    if (isRecording){
        stopRecording();
    }
    isReading = 0;
    currentValue = -1;
}

/* Start recording buffer */
void PressureSensor::startRecording()
{
    if (isRecording){ //ignore repeat requests
        return;
    }
    if (! isReading){
        startReading();
    }
    if (firstBufBlock != NULL){ //delete any old buffer
        deleteBufferBlock(firstBufBlock);
        firstBufBlock = NULL;
    }
    isRecording = 1;

}

/* Stop recording samples */
void PressureSensor::stopRecording()
{
    if (! isRecording){
        return;
    }
    //stop this thread until recording has actually stopped
    shouldStopRecording = 1;
    waitcond.wait();
}
```

```
/* The thread that actually does the polling.
 * NOTE: if isRecording is set for too long, we will eventually
 * run out of memory. This would be a most unfortunate state of
 * affairs, from which we will exit somewhat ungraciously
 * (i.e. leave it up to the compiler!!).
 */
void PressureSensor::run()
{
    bufBlock* currentBlock;
    int bufCount = 0;
    while (isReading){
        currentValue = pollSensor();
        QThread::usleep(100000); //100msec
        if (isRecording){
            if (firstBufBlock == NULL){ //it's the first loop
                currentBlock = firstBufBlock = new bufBlock;
                bufCount = 0;
            }
            if (bufCount == BLOCK_SZ){ //currentBlock is full
                currentBlock = (currentBlock->next = new bufBlock);
                bufCount = 0;
            }
            //safe to do it hear although not always needed
            currentBlock->next = NULL;
            currentBlock->dpoint[bufCount] = (char)currentValue;
            bufCount++;
            //This apparently ridiculous bit of code ensures
            //that startRecording() never tries to delete a buffer
            //that this thread is writing to; i.e. it prevents
            //a race condition in the case of a quick stopRecording()
            //startRecording() sequence. It also ensures that
            //unused bufBlock entries are zeroed.
            if (shouldStopRecording){
                //write remainder of currentBlock with 0
                for (; bufCount < BLOCK_SZ; bufCount++){
                    currentBlock->dpoint[bufCount] = 0;
                }
                shouldStopRecording = isRecording = 0;
                //any waiting threads can resume
                waitcond.wakeAll();
            }
        }
    }
}


/* Return current sample value */
int PressureSensor::getCurrentValue()
{
```

```
    return currentValue;
}

/* Recursively free memory allocated to a bufBlock  */
void PressureSensor::deleteBufferBlock(bufBlock* bb)
{
    if (bb->next != NULL){
        deleteBufferBlock(bb->next);
    }
    delete bb;
}

/* Write buffer to output stream os. */
int PressureSensor::writeBuffer(ostream& os)
{
    if (isRecording){
        stopRecording();
    }
    if (firstBufBlock != NULL){
        try { //after all, I/O is quite likely to go wrong
              // if anything does...
            writeBuffer(os, firstBufBlock);
        } catch(...) {
            return 0; //an oops occurred.
        }
    }
    return -1; //success if (a)there was nothing to do or (b)we did it.
}

/* The recursive routine that actually does the writing */
void PressureSensor::writeBuffer(ostream& os, bufBlock* bb)
{
    for (int i=0; i < BLOCK_SZ; i++){ //write buffer contents
        //cast to int so stream displays sanely
        os << (int)bb->dpoint[i] << "\t";
    }
    if (bb->next != NULL){
        //do the same for next buffer
        writeBuffer(os, bb->next);
    }
}
```

The PressureSensor class does not deal with the low-level aspects of communicating with the pressure sensor hardware. Instead, this task is delegated to a C function defined in SENSORGUTS.H and implemented in SENSORGUTS.C:

```
/* SensorGuts.h - Contains the function that actually calls the parapin
 * library. This must be separate because of the C / C++ linkage issue
 */
```

```
/* Communicate with the pressure sensor through the parapin lib */

#ifndef _sensorguts_
#define _sensorguts_

extern "C" int pollSensor();


#endif /* _sensorguts_ */

/* SensorGuts.c - Contains the implementation of pollSensor()
 * Warning: pollSensor is NOT threadsafe, and I don't even want
 * to think what the repercussions might be of randomly calling
 * it whilst is busy playing parallel ports!! Therefore, the
 * calling function _MUST_ ensure thread safety
 */

#include <parapin.h>
#include <sys/time.h>
#include <sys/types.h>
#include <unistd.h>

/* Pin definitions for ADC */
#define VCC LP_PIN01     //switchable
#define DB0 LP_PIN02    // - data range (switchable)
#define DB1 LP_PIN03    // -
#define DB2 LP_PIN04    // -
#define DB3 LP_PIN05    // -
#define DB4 LP_PIN06    // -
#define DB5 LP_PIN07    // -
#define DB6 LP_PIN08    // -
#define DB7 LP_PIN09    // - end data range
#define INTR LP_PIN15    //Non-switchable
#define WR LP_PIN16    //switchable
#define RD LP_PIN17    //switchable
#define GND LP_PIN18    //gnd
/*-- End pine definitions --*/

int pinsConfigured = 0; //we only want to set up the port first time

void waitMicrosec(int usec); //a wrapper for a microsec pause

void configPins(); //init parallel pins


/* Actually poll the sensor according to the chip specs */
int pollSensor()

#ifdef NO_PSENSOR
{
```

```
    static char counter;
    return ++counter & 127;
}
#else
{
    int val, pins, count;
    if (pinsConfigured == 0){
        configPins();
    }

    //briefly set WR low and then wait for INTR to go high
    clear_pin(WR);
    waitMicrosec(1); //200nano-sec wait required
    set_pin(WR);
    waitMicrosec(1);
    for (count = 0; pin_is_set(INTR) != 0 && count < 128; ++count){
        //conversion normally takes 114usec according to chip specs
        waitMicrosec(500);
    }
    //set RD low, read value and RD high
    clear_pin(RD);
    waitMicrosec(1); //200nano-sec wait required
    pins = pin_is_set(DB0 | DB1 | DB2 | DB3 | DB4 | DB5 | DB6 | DB7);
    set_pin(RD);

    //turn value into a sane number
    val = 0;
    if (pins & DB0) val |= 1;
    if (pins & DB1) val |= 2;
    if (pins & DB2) val |= 4;
    if (pins & DB3) val |= 8;
    if (pins & DB4) val |= 16;
    if (pins & DB5) val |= 32;
    if (pins & DB6) val |= 64;
    if (pins & DB7) val |= 128;
    return val;
}
#endif


/* Pause for a given number of microseconds */
void waitMicrosec(int usec)
{
    struct timeval timeout;
    timeout.tv_sec = 0;
    timeout.tv_usec = usec;
    select(0, NULL, NULL, NULL, &timeout);
}

/* Set the parallel port pins in the correct I/O configuration
```

```
 * and init state
 */
void configPins()
{
    //DataBus is for INPUT
    pin_input_mode(DB0 | DB1 | DB2 | DB3 | DB4 | DB5 | DB6 | DB7);
    pin_output_mode(VCC | WR | RD);
    set_pin(VCC); //for the power supply
    set_pin(WR | RD); //these should be normally high

    //INTR is for INPUT, but this is not a switchable pin, so
    //no config is needed

    //and done...
    pinsConfigured = 1;
}
```

To run this software, all the source files should be compiled and linked with Parapin and the multi-threaded version of QT; the binary file must be *setuid* as the *root* user. A *Makefile* that automates the process of compiling and linking these files and setting the *setuid* bit is included below.

```
# Makefile for Experiment Two program
# use "make" to compile and "make setuid" to set security on binary

# Define NO_PSENSOR to compile the program without support for the
# pressure sensor. The motivation for this option is to allow
# parts of the program that are not dependent on the pressure sensor
# itself to be worked on without any special hardware, or fear of
# frying the computer! When NO_SENSOR is defined, the pollSensor()
# function defined in SensorGuts.h returns 32, and the parapin library
# is left uninitialised

EXTRA_CFLAGS = -DNO_PSENSOR

CFLAGS = -DQT_THREAD_SUPPORT -I/usr/local/qt/include \
         -O -Wall $(EXTRA_CFLAGS)
LDFLAGS = -L/usr/local/qt/lib -lqt-mt -lparapin

BINARY = ../bin/exp2
OBJS =      \
    Experiment.o \
    ExpDisplay.o \
    PressureSensor.o \
    moc_PressureSensor.o \
    moc_Experiment.o \
    moc_ExpDisplay.o \
    sensorguts.o \
```

```
    qtmain.o \
    main.o

all: $(BINARY)

clean:
    rm -f $(OBJS)
    rm -f $(BINARY)

setuid:
    chown root $(BINARY)
    chmod a+s $(BINARY)

#Make exp2 binary by linking all the object files
$(BINARY): $(OBJS)
    g++ -o $(BINARY) $(OBJS) $(LDFLAGS)


# Create moc files for all headers
# Note: it's fine to do it for all, as non-QT files are ignored
moc_%.cpp: %.h
    moc -o $@ $<

# Standard compilation of C++ sources
%.o: %.cpp
    g++ -c -o $@ $(CFLAGS) $<

# Standard compilation of C sources
%.o: %.c
    gcc -c -o $@ $(CFLAGS) $<
```

After successful compilation, issue the command "exp2 <filename>" where "<filename>" is the name of a file to which the results of an experiment session will be written. For further information, please contact the author of this thesis.

In addition to the software written to run the experiment, a second program—written in Java—was used to convert the raw data from the experiment sessions into a form in which it could be analysed. This included performing the various mathematical transformations of the pressure sensor traces (described in Chapter Eight). The program consists of a set routines for parsing the data files and a set of mathematical functions to operate on the pressure data. Its behaviour can be controlled by changing the value NUM_SAMPLES (which defines how many values the program should output when downsampling), and by chaining together combinations of the mathematical functions. Examples are provided in the code itself. Output from the program consists of a single list of data points, optionally normalized per participant, in a form suitable for analysing within SPSS.

```
package analysis;
import java.io.*;
import java.util.*;

public class DataFormatter {
    public static final String datadir =
       "/home/mml1000/phd-working/phd_full/experiments/exp_2/raw-data/";
    public static final String textdir =
       "/home/mml1000/phd-working/phd_full/experiments/exp_2/texts/";
    public static List propositions,factsA,factsB,isFactsATrue;
    public static Map sevenPtRatings;

    //number of propositions/fact/etc
    public static final int NUM_PROPS = 20;
    //number to represent missing values
    public static final int SILLY_NUMBER = -9999;
    //number of points to which to downsample
    public static final int NUM_SAMPLES = 6;
    public static final String outFileName =
       "/tmp/analysis_out.txt";
    public static final String[] dataFiles = {
            "abb27.A.txt", "ajs94.A.txt", "ali.B.txt", "apt.A.txt",
            "ashley.A.txt", "csr21.B.txt", "cvw21.A.txt", "jhd25.A.txt",
            "marcc.B.txt", "markd.B.txt", "mkb.A.txt", "owen.B.txt",
            "rachel.B.txt", "sam.B.txt", "timh.A.txt", "mk270.A.txt",
            "paul.B.txt"
    };

    public static void main(String[] args) throws IOException
    {
        //initialise databases
        propositions = readLines(textdir + "propositions.txt");
        factsA = readLines(textdir + "factsA.txt");
        //there's an error in factsA, which means Captain Scott's death
        //is re-enacted. Deal with this here
        factsA.remove(19);
        factsB = readLines(textdir + "factsB.txt");
        isFactsATrue = readLines(textdir + "isFactACorrectAnswer.txt");

        //get 7ptdata into hashmap
        sevenPtRatings = createRatingsMap();

        //open output file for results & write header
        PrintWriter output = new PrintWriter
                             (new FileWriter(outFileName), false);
        String hdr = "Person\tTrial\tPlayOrder\tRating\t";
        for(int i=0; i<NUM_SAMPLES;++i) hdr += "Sample" + (i+1) + "\t";
```

```
hdr += "RespTime";
output.println(hdr);


//process data files
for(int df=0; df<dataFiles.length; ++df){
    List answers = createListOfCorrectAnswers
                      (readRawData(dataFiles[df]));
    double[][] trialsOut = new double[answers.size()][];
    for(int trial=0; trial<answers.size();++trial){
        double[] pressures = cutTrailingZeros
                (((Answer)answers.get(trial)).pressures);

        /* Decide which procedures to carry out here. This is
         * accomplished by  invoking some combination of the
         * mathematical routines in a sensible order.
         * double[] independents must end up containing the
         * resulting data set. Some sensible combinations of
         * functions whose output could be assigned to it are
         * listed below:
         *  = downSample(pressures,NUM_SAMPLES);
         *  = downSample(absolute
         *          (firstDerivative(pressures,5)),NUM_SAMPLES);
         * = downSample(absolute(firstDerivative(firstDerivative
         *              (pressures,5),5)),NUM_SAMPLES);
         * = downSampleVariance(pressures,NUM_SAMPLES);
         * = downSampleVariance(firstDerivative(pressures,5),
         *               NUM_SAMPLES);
         * = downSampleVariance(absolute(firstDerivative
         *                          (pressures,5)),NUM_SAMPLES);
         * = downSampleVariance(absolute(firstDerivative
         *          (firstDerivative(pressures,5),5)),NUM_SAMPLES);
         */

        double[] independents = downSample(pressures,NUM_SAMPLES);

        //for each trial we have playorder, rating,
        //independents, a responsetime
        Answer ans = ((Answer)answers.get(trial));
        trialsOut[trial] = new double[independents.length + 3];
        trialsOut[trial][0] = ans.playOrder;
        trialsOut[trial][1] = ans.rating;
        System.arraycopy(independents,0,trialsOut[trial],
                                     2,independents.length);
        trialsOut[trial][independents.length+2] = ans.respTime;

    }
    //now perform normalization on each COLUMN of numbers
    //except playorder
   for(int col=1; col<trialsOut[0].length; col++){
```

```
                //create an array to normalize
                double[] colData = new double[trialsOut.length];
                for(int row=0; row<trialsOut.length; row++){
                    colData[row] = trialsOut[row][col];
                }
                //remove outrageously outlying response times
                if(col == trialsOut[0].length - 1)
                             markSilly(colData,1000,20000);
                //normalize
                normalizeArray(colData);
                //put the normalized data back
                for(int row=0; row<trialsOut.length; row++){
                    trialsOut[row][col] = colData[row];
                }
            }

            //finally, print the stuff out!
            for(int row=0; row<trialsOut.length;row++){
                output.print(df + "\t" + row + "\t");
                for(int col=0; col<trialsOut[row].length; col++){
                    output.print(trialsOut[row][col] + "\t");
                }
                output.print("\n");
            }
        }

        //close output file
        output.close();
    }

    /* return the first derivative of a passed array */
    public static double[] firstDerivative(double[] data, int lag)
    {
        if(data.length < 1) return data;
        double[] retval = new double[data.length - lag];
        for(int i=0;i<retval.length;i++){
            retval[i] = data[i+lag] - data[i];
        }
        return retval;
    }

    /* change an array of doubles into absolute values*/
    public static double[] absolute(double[] data)
    {
        for(int i=0; i<data.length; i++){
                data[i] = Math.abs(data[i]);
        }
        return data;
    }
```

```
/* calculate a set of numSample points that represent the variance
 * of the data as downsampled by a triangle-shaped window function
 */
public static double[] downSampleVariance(double[] data,
                                          int numSamples)
{
    double newdata[] = new double[numSamples];
    //if no data, return a set of -1 (we never have -ve in real data)
    if(data.length == 0){
        for(int i=0;i<numSamples;i++) newdata[i] = SILLY_NUMBER;
        return newdata;
    }

    //we have data, so downsample
    for(int i=0;i<numSamples;i++){
        double sum = 0;
        double sumSquares = 0;
        double sumWeights = 0;
        for(int j=0;j<data.length;j++){
            double x = (double)j/(double)data.length
                     * (double)(numSamples+1) - i;
            if(x>1.0) x = 2.0 - x; //x moves between 0 and 1 and back
            if(x<0.0) x = 0.0;     //
            sum += x * data[j];
            sumSquares += x * data[j] * data[j];
            sumWeights += x;
        }
        double weightedMean = sum / sumWeights;
        newdata[i] = (sumSquares/sumWeights - weightedMean
                        * weightedMean);
    }
    return newdata;
}


/* remove trailing zeros */
public static double[] cutTrailingZeros(double[] data)
{
    int len = 0;
    for(int i=0; i<data.length;++i){
        if(data[i] != 0) len = i+1;
    }
    double retval[] = new double[len];
    System.arraycopy(data,0,retval,0,len);
    return retval;
}

/* calculate a set of numSample points that represent the data
   as downsampled by a triangle-shaped window function */
```

```
public static double[] downSample(double[] data, int numSamples)
{
    double[] newdata = new double[numSamples];

    //if there is no data, just return a set of -1
    //(we never have -ve in real data)
    if(data.length == 0){
        for(int i=0;i<numSamples;i++) newdata[i] = SILLY_NUMBER;
        return newdata;
    }

    //we have data, so do downsample
    for(int i=0; i<numSamples; ++i){
        double rt=0; //running total
        //work out an integration of the area under i trianguli
        for(int j=0; j<data.length; ++j){
            //for each triangle, x is 0 at left base of triangle
            //1 at peak and 2 at right base
            double x = (double)j/(double)data.length
                     * (double)(numSamples+1) - i;
            if(x > 0.0 && x < 1.0) rt += x*data[j];
            if(x >= 1.0 && x < 2.0) rt += (2-x)*data[j];
        }
        newdata[i] = rt/data.length * numSamples;
    }
    return newdata;
}


/* Replace any numbers that fall outside of the range min < x < max
   with SILLY_NUMBER */
public static void markSilly(double[] data, int min, int max)
{
    for(int i=0; i<data.length; i++){
        if(data[i] <= min || data[i] >= max){
            data[i] = SILLY_NUMBER;
        }
    }
}

/* generic normalization routine */
public static void normalizeArray(double[] data)
{
    //find mean and stdev
    double sum = 0, sumOfSquares = 0;
    int validNums = 0;
    for(int i=0;i<data.length;i++){
        //don't include SILLY_NUMBER values
        if(data[i] != SILLY_NUMBER){
            sum += data[i];
```

```
                sumOfSquares += data[i]*data[i];
                validNums++;
            }
        }
        double mean = sum/validNums;
        double stdev = Math.sqrt((sumOfSquares - sum*mean)
                     / (validNums - 1));

        //normalise, but leave SILLY_NUMBERs alone
        for(int i=0;i<data.length;i++){
            if(data[i] != SILLY_NUMBER){
                data[i] -= mean;
                data[i] /= stdev;
            }
        }
    }


    /* Create a list of correct answers */
    public static List createListOfCorrectAnswers(Answer[] origAns)
    {
        List lst = new ArrayList();
        for(int i=0; i<NUM_PROPS; ++i){
            if(origAns[i].isCorrect){
                lst.add(origAns[i]);
            }
        }
        return lst;
    }

    /* Read raw data from a results file and put data into sane form */
    public static Answer[] readRawData(String filename)
       throws IOException
    {
        Answer[] answers = new Answer[NUM_PROPS];

        //init array of Answers
        for(int i=0; i<NUM_PROPS; ++i){
            answers[i] = new Answer();
        }

       //crudely check filename for A or B-ness
       boolean usesFactsA = "A.txt".equals
                    (filename.substring(filename.length() - 5));

        //fill up array by parsing file (first part of experiment)
        BufferedReader br = new BufferedReader
                    (new FileReader(datadir + filename));
        String line;
        br.readLine(); //discard two lines
```

```
    br.readLine();
    //printDatabase();
    for(int i=0; i<NUM_PROPS; ++i){
        line = br.readLine();
        if(!line.substring(0,13).equals("Proposition: ")){
            oops("Ouch!! Proposition " + i +" expected.");
        }

        int whichProp = propositions.indexOf(line.substring(13));
        if(whichProp < 0){
            oops("Proposition \""+line.substring(13)+"\" not found.");
        }
        Answer ans = answers[whichProp];
        br.readLine(); // Skip music line
        if(!br.readLine().equals("Sensor Data: ")){
            oops("Ouch!! Sensor data line oopsed for " + i);
        }

        //read in sensor data
        StringTokenizer sensData = new StringTokenizer
                          (br.readLine(),"\t",false);
        ans.pressures = new double[sensData.countTokens()];
        for(int j=0; sensData.hasMoreTokens(); j++){
            ans.pressures[j] = Integer.parseInt(sensData.nextToken());
        }
        //skip blank line
        br.readLine();

        // add playOrder and 7pt scale rating
        ans.playOrder = i;
        ans.rating = getRating(filename,i);
    }

    //parse second part of file
    br.readLine();
    //loop 21 times thanks to Captain Scott
for(int i=0; i<(usesFactsA ? NUM_PROPS+1 : NUM_PROPS); ++i){
         line = br.readLine();
         StringTokenizer quesData = new StringTokenizer
                                          (line,"\t",false);
        if(quesData.countTokens() != 3)
                oops("Malformed question line: \"" + line + "\"");
        String question = quesData.nextToken();
        if(usesFactsA && (question.equals(factsA.get(19))
                        || question.equals(factsB.get(19)))){
            //we have a Facts A Captain Scott situation
            Answer ans = answers[19];
            if(ans.respTime == 0.0){
                //we have not yet encountered the Captain
                ans.respTime = Integer.parseInt(quesData.nextToken());
```

```
                boolean correctAnswer = !question
                                       .equals(factsA.get(19))
                                       ^ isFactsATrue.get(19)
                                       .equals("true");
                // yes indeed. It would appear that my
                //Experiment running program had what
                // is known in the trade as a "bug".
                //The button labelled "True" returned "False"
                // and vice versa. Fixed here!
                boolean actualAnswer = !quesData.nextToken()
                                        .equals("True");
                ans.isCorrect = (actualAnswer == correctAnswer);
            }
        } else {
            int whichQues = (usesFactsA ? factsA : factsB)
                          .indexOf(question);
            if(whichQues < 0)
                oops("I cannot find question \"" + line + "\"");
            Answer ans = answers[whichQues];
            ans.respTime = Integer.parseInt(quesData.nextToken());
            //returns the correctness of the answer
            boolean correctAnswer = !usesFactsA ^ isFactsATrue
                                .get(whichQues).equals("true");
            //another bugfix for the true-false problem
            boolean actualAnswer = !quesData.nextToken()
                                              .equals("True");
            ans.isCorrect =  (actualAnswer == correctAnswer);
        }

     }
    if(!br.readLine().equals("__END OF SESSION__"))
       oops("We have not reached the end of file as expected");
    return answers;
}

/* find the 7pt rating associated with a pressure trace */
public static int getRating(String filename, int trial)
{
    int[] ratings = (int[])sevenPtRatings.get(filename);
    //is there a rating sheet for his participant?
    if(ratings == null){
        return SILLY_NUMBER;
    }
    return ratings[trial];
}

/* Reads 7pt ratings data from file into a map
   that maps pressure sensor data filename to an array of ratings */
public static Map createRatingsMap() throws IOException
{
```

```
        Map ratings = new HashMap();
        Iterator it = readLines(datadir + "ratings.txt").iterator();

        //go through file one line at a time, splitting it
        //into name and int[]
        while(it.hasNext()){
            StringTokenizer stk =
                  new StringTokenizer((String)it.next(),",",false);

            //there should be 19 tokens. If not, pretend doesn't exist
            if(stk.countTokens()==19){
                String key = stk.nextToken();
                //two dummy entries for practice trials
                int[] vals = new int[20];
                vals[0] = vals[1] = 0;
                for(int i=2;stk.hasMoreTokens();i++){
                    vals[i] = Integer.parseInt(stk.nextToken());
                }
                ratings.put(key,vals);
            }
        }
        return ratings;
    }

    public static List readLines(String filename) throws IOException
    {
        BufferedReader br = new BufferedReader(new FileReader(filename));
        List lst = new ArrayList();
        String line;
        while((line = br.readLine()) != null){
            lst.add(line);
        }
        return lst;
    }


    public static void oops(String theOops)
    {
        System.out.println(theOops);
        System.exit(-52); //after all, why not!
    }

}

class Answer
{
    public double[] pressures;
    public double respTime;
    public boolean isCorrect;
    public int playOrder; //order in which pt1 trial was played
```

```
    public int rating; //from rating scale input file
    public String toString()
    {
        String pp="{";
        if (pressures != null){
            if (pp.length() != 0) pp += pressures[0];
            for(int i=1; i< pressures.length; ++i){
                pp += ", " + pressures[i];
            }
        }
        pp += "}";

        return getClass().getName() +
               " [respTime=" +  respTime +
               ", isCorrect=" + isCorrect +
               ", playOrder=" + playOrder +
               ", rating=" + rating +
               ", Pressure=" + pp + "]";
    }
}
```

The Pressure Sensor

The Pressure Sensor—used in the experiments presented in Chapters Eight and Nine—converts pressure applied to a hand-held unit into a number between 0 and 127; the data is read and the operation controlled via a PC parallel port.[1] It consists of a pressure transducer (Honeywell FSG15N1A) mounted in a hand-held box, an 8-bit analogue-to-digital convertor (ADC080X type) and circuitry to connect the two. The circuit was designed and built by the author of this thesis and Alistair Turnbull (Computer Laboratory, University of Cambridge).

The circuit diagram is presented in figure A.1. The Wheatstone bridge arrangement on the left-hand side is the pressure transducer itself, which provides a differential voltage in the milli-volt range. A pair of buffer amplifiers (in an LM747 chip) and related circuitry magnifies and anchors the differential so that it can provides a range suitable for driving the A-D convertor. A schematic representation of the ADC080X convertor is presented at the right-hand side of the circuit: analogue inputs are shown on the left, the clock on top, power supply and the data bus on the right, and control lines on the bottom. The numbers on the far right of the diagram show the pins on the parallel port to which the convertor's I/O lines

[1]The intention was to have a range of between 0 and 255, but a limitation inherent in the design of the analogue circuit resulted in the narrower range.

are connected. The circuit runs at +5v DC. In our implementation, it was powered by a 9 volt battery, regulated down to 5v by a KA78LXX (shown at the bottom left-hand side of the diagram). For further information about how to build and write software that can use this circuit, please contact the author of this thesis.

Software used in Chapter Nine

The experiment reported in Chapter Nine was controlled by an adapted version of that used in the experiment reported in Chapter Eight. All of the Pressure Sensor code (in PRESSURESENSOR.H, PRESSURESENSOR.CPP, SENSORGUTS.H AND SENSORGUTS.C) remains unaltered in the new version. Likewise, the initialisation code MAIN.C and QTMAIN.CPP remains essentially identical, except for one tiny change that allows two command-line parameters to be passed to the Experiment constructor as opposed to one; the second parameter determines whether the experiment session should present music and pictures or just music alone. By contrast, the main Experiment class (EXPERIMENT.H and EXPERIMENT.CPP) was re-written almost from scratch to cater for very different control flow and data structures in the Chapter Nine experiment. The new class and its implementation are printed below:

```
/* Experiment - class that encapsulates the experiment.
 * controls sesssion flow, stimuli, responses, etc */

#ifndef _Experiment_
#define _Experiment_

#include <qobject.h>
#include <qtimer.h>
#include <qstring.h>
#include <qevent.h>
#include <fstream.h>
#include <iostream.h>

#include "PressureSensor.h"
#include "PicDisplay.h"
#include "SensorDisplay.h"
#include "RateDisplay.h"
#include "InstructDisplay.h"

class Experiment : public QObject,QThread
{
    Q_OBJECT
```

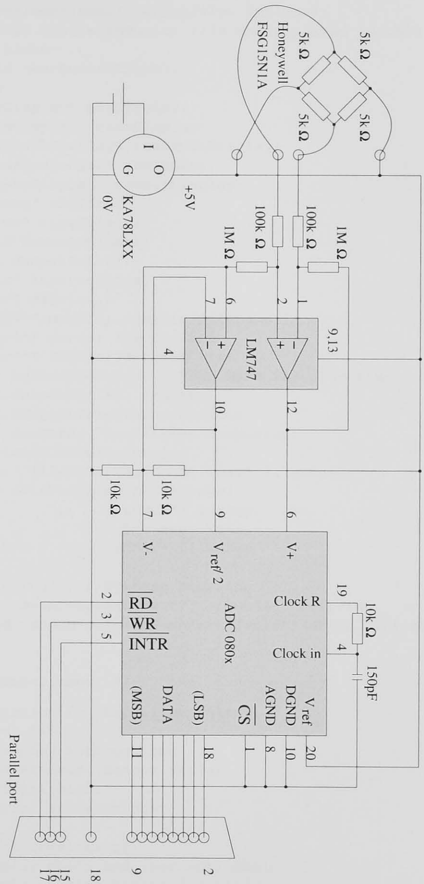

Figure A.1: Circuit diagram for Pressure Sensor

```
    public:
        Experiment(char* outputfile, int mus);
        static bool withMusic; //is music played in this session
    public slots:
        void hearButtonClick();
    private:
        PicDisplay* picDisplay;
        RateDisplay* rateDisplay;
        InstructDisplay* instructDisplay;
        PressureSensor* theSensor;
        SensorDisplay* sensorDisplay;
        QTimer* theTimer;
        QTime* timeObject;
        void doNextThing();
        int doNextTrial();
        char* instructions;
        char* debrief;
        char*** musics; //music filename[]
        char*** words; //word list
        char*** trials; //trial list
        int currentSection; //keep tabs for play time
        int currentTrial; //ditto
        int currentTrialPart; //ditto
        int numWords, numTrials, numMusics;
        ofstream resultsStream;
        char* fileToString(const char* filename);
        int fileToStringTokenArrays(
                        char*** &array,
                        int numTokens,
                        char* filename
        );
        void run(); //QThread override
        bool event(QEvent* evt); //QObject override
        void randomiseKnownSizeArray(void** array, int size);
};

#endif /* _Experiment_ */

/* Implementation for Experiment class.
 * Experiment flow:
 *    (1) Display instructions
 *    (2) Do Pressure Sensor Feeler
 *  (5) Do Trial block
 *  (6) Debrief Instructions
 *
 * Each Trial consists of:
 *  (1) Display photo and (perhaps) music
 *  (2) Display Rating Scales for titles and words
 *  (3) Write out pressur data (perhaps), and ratings
 *  (4) Display please wait for next trial (n secs)
 *
```

```
 * The order of title display and word display permuted.
 * Trial order randomised; music randomly assigned.
 *
 */

#include "Experiment.h"
#include <qapplication.h>
#include <sys/stat.h>
#include <unistd.h>
#include <iomanip.h>
#include <qdatetime.h>
#include <string.h>
#include <stdlib.h>
#include <time.h>
#include <stdio.h>

/* defines for filenames */
#define TRIALFILE "../stimuli/trials.txt" //trial file
#define MUSICFILE "../stimuli/musics.txt" //music extract FILENAMES
#define WORDFILE "../stimuli/words.txt" //list ofwords
#define INSTRUCTFILEM "../stimuli/instruct-mus.txt" //instructions
#define INSTRUCTFILENM "../stimuli/instruct-nm.txt"
#define DEBRIEFFILE "../stimuli/debrief.txt" //debrief

#define SOUND_PLAYER "/usr/bin/mpg123 " //NB the space!
#define PICSHOWTIME 30000 //3secs
#define TITLES 4

bool Experiment::withMusic = 0;
Experiment::Experiment(char* outputfile, int mus)
{
    withMusic = (mus != 0);
    currentSection = currentTrial = currentTrialPart = 0;
    theSensor = new PressureSensor();

    //import instructions and debrief
    if (    (instructions = fileToString(
                withMusic ? INSTRUCTFILEM : INSTRUCTFILENM
            )) == NULL ||
            (debrief = fileToString(DEBRIEFFILE)) == NULL
        ){
             qFatal("One of the instruction files is broken.\n");
    }

    //read in music, words and trials
    numMusics = fileToStringTokenArrays(musics, 1, MUSICFILE)-1;
    numTrials = fileToStringTokenArrays(trials, TITLES+1, TRIALFILE)-
1;
    numWords = fileToStringTokenArrays(words, 1, WORDFILE)-1;
```

```
//did it work?
if (musics==NULL) qFatal("Music file broken.\n");
if (words==NULL) qFatal("Word file broken.\n");
if (trials==NULL) qFatal("Trial file broken.\n");
if (numTrials != numMusics){
    qFatal("Non-matching numbers of extracts and trials.\n");
}

//check music files are all statable
struct stat fi;
for (int i=0; musics[i] != NULL; i++){
    if (stat(musics[i][0],&fi)){
        qFatal("Music file %s is broken.\n", musics[i][0]);
    }
}

//set up random seed
srand((unsigned int)time(NULL));

//create random presentation order for musics and trials
randomiseKnownSizeArray((void**)musics, numMusics);
randomiseKnownSizeArray((void**)trials, numTrials);

//open a resultsfile for writing
resultsStream.open(outputfile, ios::out|ios::noreplace|ios::stdio);
if (resultsStream.bad()){
    qFatal("Could not open %s for writing.\n", outputfile);
}
resultsStream
        << "Experiment Session started on "
        << QDateTime::currentDateTime().toString().latin1()
        << "\n";

//initialise environment
picDisplay = new PicDisplay(NULL, "Pictures");
instructDisplay = new InstructDisplay(NULL, "Instructs");
sensorDisplay = new SensorDisplay(NULL, "SensorDisp");
theTimer = new QTimer(picDisplay);
rateDisplay = new RateDisplay(NULL, "Ratings", TITLES, numWords);

connect( //instructions read
   instructDisplay, SIGNAL(buttonClicked()),
   this, SLOT(hearButtonClick())
);
connect( //timer timeout
    theTimer, SIGNAL(timeout()),
    this, SLOT(hearButtonClick())
);
connect( //sensor test over
    sensorDisplay, SIGNAL(buttonClicked()),
```

```
        this, SLOT(hearButtonClick())
    );

    connect( //ratings done
        rateDisplay, SIGNAL(buttonClicked()),
        this, SLOT(hearButtonClick())
    );

    qApp->setMainWidget(picDisplay);

    doNextThing();

}

/* SLOT */
void Experiment::hearButtonClick()
{
    doNextThing();
}

/* The main experiment loop. Control branches from here depending
 * on where we are in the experiment.
 */
void Experiment::doNextThing()
{
    switch (currentSection){
        case 0: //initial instructions
            currentSection++;
            instructDisplay->showInstructions
                (instructions, "Click here to start experiment");
        break;

        case 1: //sensor testing
            currentSection++;
            if (withMusic){
                sensorDisplay->displayMe(theSensor);
            } else {
                doNextThing();
            }
        break;

        case 2: //trials
            if (doNextTrial() == 0){
                currentSection++;
                doNextThing();
            }
        break;

        case 3: //debrief instructions
            currentSection++;
```

```
            instructDisplay->showInstructions
                    (debrief, "Click here to finish.");
        break;

        default: //experiment has ended
            resultsStream << "__END OF SESSION__";
            resultsStream.flush();
            resultsStream.close();
            qApp->exit(0);
    }
}


/* Runs the next question, based on currentTrial.
 * Returns -1 if there is something to do.
 * Once there are no more questions to do, returns 0
 */
int Experiment::doNextTrial()
{
    //finish
    if (trials[currentTrial] == NULL){
        return 0;
    }

    switch (currentTrialPart){
        case 0: //time to display picture & maybe music
            currentTrialPart++;

            //write out trial info
            resultsStream << "Trial " << currentTrial
                          << "\nPicture="
                          << trials[currentTrial][0]
                          << "\n\n";

            picDisplay->showPicture(
                trials[currentTrial][0]
            );
            if (withMusic){
                //play music and sensor in background thread
                start();
            } else {
                //trigger doNextThing() after PICSHOWTIMEmillis
                theTimer->start(PICSHOWTIME, TRUE);
            }
        break;
        case 1: //time to display rating scales
            currentTrialPart++;
            picDisplay->hide();
            randomiseKnownSizeArray(
                (void**)&trials[currentTrial][1],TITLES
```

```
            );
            randomiseKnownSizeArray((void**)words, numWords);
            rateDisplay->displayMe(
                &trials[currentTrial][1],
                words
            );
        break;
        case 2: //time to record results and move on
            rateDisplay->writeResults(resultsStream);
            currentTrialPart = 0;
            currentTrial++;
            doNextThing();
        break;
        default:
            qFatal("Something utterly weird has happened.\n");
    }
    return -1;
}


/* Start the sensor, play the music and write sensor data */
void Experiment::run()
{
        //prepare commandline for media player
        char* cmd = new char[
                sizeof(SOUND_PLAYER) +
                strlen(musics[currentTrial][0]) +
                sizeof(" 2>/dev/null")
        ];
        strcpy(cmd, SOUND_PLAYER);
        strcat(cmd, musics[currentTrial][0]);
        strcat(cmd, " 2>/dev/null");

        //record && play
        theSensor->startRecording();
        system(cmd);
        QThread::msleep(500);
        theSensor->stopRecording();
        delete[] cmd;
        cmd = NULL;

        //writeout data
        resultsStream << "Music: "
                      << musics[currentTrial][0]
                      << "\n";
        resultsStream << "Sensor Data: \n";
        theSensor->writeBuffer(resultsStream);
        resultsStream << "\n";

        //post event to say we're done, so control flow
```

```
        //can continue in the main thread
        QThread::postEvent(this, new QCustomEvent(10001));
}

/* OVERRIDE QObject::event() to handle finishing music */
bool Experiment::event(QEvent* evt)
{
    if (evt->type() == 10001){
        doNextThing();
        return true;
    } else {
        return QObject::event(evt);
    }
}

/* Grab the contents of a file and put them into a single
 * null-terminated string; returns pointer to the newly
 * created string or null if it fails*/
char* Experiment::fileToString(const char* filename)
{
    //allocate requisite memory
    struct stat fi;
    if (stat(filename,&fi) != 0){
        return NULL;
    }
    char* buffer = new char[fi.st_size + 1];

    //open file
    ifstream ifs(filename, ios::in);
    if (ifs.bad()){
        return NULL;
    }
    ifs.unsetf(ios::skipws); //do not ignore '\n' and ' '

    //fill buffer && return
    for (int i = 0; i < fi.st_size; i++){
        ifs >> buffer[i];
    }
    buffer[fi.st_size] = '\0';
    ifs.close();
    return buffer;
}

/* Grab the contents of a file and put them into string array arrays,
 * one string array per line of the original file (i.e. split on '\n').
 * one string per token (i.e. split on ',').
 * Memory is allocated as required; final element of outer array is
 * NULL so the end can be determined. Returns the number of elements
 * in the array (_INCLUDING_ the final NULL element), or in case of
 * problem, it just returns 0 and sets char*** array to NULL.
```

```
 * Token arrays (i.e. the inner arrays) are NOT NULL-delimited, but if
 * any line has the wrong number of tokens, the array returns null
 */
int Experiment::fileToStringTokenArrays
    (char*** &array, int numTokens ,char* filename)
{
    //grab to temporary buffer
    char* tmp;
    if ((tmp = fileToString(filename)) == NULL){
        array = NULL;
        return 0;
    }

    //work out no. lines required, and allocate memory for pointers
    int numlines = 0;
    for (int i=0; tmp[i] != '\0'; i++){
        if (tmp[i] == '\n') numlines++;
    }
    char*** tmpArray = new char**[numlines + 1];
    for (int i=0; i<numlines; i++){
        tmpArray[i] = new char*[numTokens];
    }

    //allocate the strings (one per line) and populate
    int whichln = 0; //current line
    int whichtok = 0;  //current token
    char* s = tmp; //pointer to start of it
    for (int i=0; tmp[i] != '\0'; i++){
        //at the end of line ...
        if (tmp[i] == '\n'){
            tmp[i] = '\0'; //replace '\n' with '\0'
            tmpArray[whichln][0] = s; //save pointer to start
            s = &tmp[i+1]; //s points to start of next string
            if (whichtok != numTokens-1){
                //we have not got the right number of tokens
                for (int q=0;q<numlines;q++) delete tmpArray[q];
                delete[] tmpArray;
                delete tmp;
                array = NULL;
                return 0;
            }
            whichln++;
            whichtok=0;
        }
        //at the end of token save pointer to start of
        //next one
        if (tmp[i] == ',' && whichtok < numTokens-1){
            tmp[i] = '\0';
            tmpArray[whichln][++whichtok] = &tmp[i+1];
        }
```

```
    }
    tmpArray[numlines] = NULL;
    array = tmpArray;
    return numlines+1;
}

/* Returns a pointer to an array of pointers. This array has
 * one entry for each number between 0 and size-1. They are
 * in a random order.
 */
void Experiment::randomiseKnownSizeArray(void** array, int size)
{
    void* tmp;
    int rnd;
    for (int i=0; i<size; i++){
        rnd = (int)((size*(double)rand())/(RAND_MAX+1.0));
        tmp = array[rnd];
        array[rnd] = array[i];
        array[i] = tmp;
    }
}
```

Owing to the need to display high-quality full-screen photographs and the relatively complex GUI requirements of the rating scales, the old Experiment Display class (defined in EXPDISPLAY.H and implemented in EXPDISPLAY.CPP) is replaced in this program with a set of six separate classes. InstructDisplay (defined in INSTRUCTDISPLAY.H and implemented in INSTRUCTDISPLAY.CPP) displays the experiment instructions and debriefing message. SensorDisplay (defined in SENSORDISPLAY.H and implemented in SENSORDISPLAY.CPP) provides a graphical dial for the Pressure Sensor and was used in the experiment to "train" participants in the sensor's use. PicDisplay (defined in PICDISPLAY.H and implemented in PICDISPLAY.CPP) loads a JPEG file from disk and displays it in a full-screen window. The code for these three classes is listed below:

```
/* InstructDisplay.h - The instruction display window */

#ifndef _InstructDisplay_
#define _InstructDisplay_

#include <qvbox.h>
#include <qtextview.h>
#include <qpushbutton.h>

class InstructDisplay : public QVBox
{
```

```
    Q_OBJECT
    public:
        InstructDisplay(QWidget* parent, const char* name);
        void showInstructions(char* instructions, char* buttontxt);
    signals:
        void buttonClicked();
    private slots:
        void hearButtonClick();
    private:
        QTextView* longText;
        QPushButton* button;
};

#endif /* _InstructDisplay_ */

/* InstructDisplay.cpp - implementation for InstructDisplay class
 * the display widget for Experiment Two */

#include "InstructDisplay.h"
#include <qfont.h>
#include <qcolor.h>
#include <qpalette.h>
#include <qtimer.h>
#include <qpushbutton.h>
#include <qtextview.h>
#include <qscrollview.h>

#define WND_WIDTH 800
#define WND_HEIGHT 500

InstructDisplay::InstructDisplay(QWidget *parent, const char *name)
        : QVBox (parent, name)
{
    setMinimumSize(WND_WIDTH,WND_HEIGHT);
    longText = new QTextView(this);
    longText->setHScrollBarMode(QScrollView::AlwaysOff);
    QFont* buttonFont = new QFont("Helvetica", 18, QFont::Bold);

    button = new QPushButton("Button", this, "button");
    button->setFont(*buttonFont);

    connect(
        button, SIGNAL(clicked()),
        this, SLOT(hearButtonClick())
    );
}


/* set instruction text (and button text) for display */
void InstructDisplay::showInstructions(
    char* instructions,
```

```
    char* buttontxt
)
{
    longText->setText(instructions);
    button->setText(buttontxt);
    this->show();
}

/* SLOT */
void InstructDisplay::hearButtonClick()
{
    this->hide();
    emit buttonClicked();
}

/* SensorDisplay.h - The display window for Experiment Two */

#ifndef _SensorDisplay_
#define _SensorDisplay_

#include <qvbox.h>
#include <qtimer.h>
#include <qdial.h>
#include "PressureSensor.h"

class SensorDisplay : public QVBox
{
    Q_OBJECT
    public:
        SensorDisplay(QWidget* parent, const char* name);
        void displayMe(PressureSensor* theSensor);
    signals:
        void buttonClicked();
    private slots:
        void hearButtonClick();
        void timerOut();
    private:
        QTimer* timer;
        QDial* dial;
        PressureSensor* theSensor;
};

#endif /* _SensorDisplay_ */

/* SensorDisplay.cpp - implementation for SensorDisplay class */

#include "SensorDisplay.h"
#include <qpushbutton.h>

SensorDisplay::SensorDisplay(QWidget *parent, const char *name)
        : QVBox (parent, name)
```

```
{
    dial = new QDial(this);
    dial->setMinValue(0);
    dial->setMaxValue(127);
    dial->setEnabled(FALSE);
    dial->setValue(0);
    dial->setMinimumSize(400,400);
    QPushButton* button = new QPushButton(
            "Click here to start experiment", this, "button"
    );
    timer = new QTimer(this);

    connect( //end of testing
        button, SIGNAL(clicked()),
        this, SLOT(hearButtonClick())
    );

    connect( //timer timeout
        timer, SIGNAL(timeout()),
        this, SLOT(timerOut())
    );
}

void SensorDisplay::displayMe(PressureSensor* sens)
{
    theSensor = sens;
    theSensor->startReading();
    timer->start(150,FALSE);
    this->show();
}

/* SLOT */
void SensorDisplay::timerOut()
{
    dial->setValue(theSensor->getCurrentValue());
}

/* SLOT */
void SensorDisplay::hearButtonClick()
{
    theSensor->stopReading();
    timer->stop();
    this->hide();
    emit buttonClicked();
}

/* PicDisplay.h - The display window for pictures */

#ifndef _PicDisplay_
#define _PicDisplay_
```

```
#include <qlabel.h>

class PicDisplay : public QLabel
{
    Q_OBJECT
    public:
        PicDisplay(QWidget* parent,const char* name);
        void showPicture(const char* filename);
};

#endif /* _PicDisplay_ */

/* PicDisplay.cpp - implementation for PicDisplay class */

#include "PicDisplay.h"
#include <qpixmap.h>
PicDisplay::PicDisplay(QWidget* parent, const char* name)
        : QLabel(parent, name){}

void PicDisplay::showPicture(const char* filename)
{
    QPixmap* piccie = new QPixmap(filename);
    this->setPixmap(*piccie);
    this->setScaledContents(TRUE);
    this->showFullScreen();
}
```

The other three GUI classes—RateDisplay, RatingScale and SaneSlider—are all concerned with the window that presents the various titles and words and asks for ratings. SaneSlider (SANESLIDER.H and SANESLIDER.CPP) is a subclass of the standard Qt Slider widget, configured to behave as a discrete seven-point scale; RatingScale (RATINGSCALE.H and RATINGSCALE.CPP) encapsulates the SaneSlider and a text label (ie. it contains all the code needed to present and record a single rating; RateDisplay (RATEDISPLAY.H and RATEDISPLAY.CPP) is responsible for building a window, filling it with the relevant RatingScales for each trial, and writing the results out to a file. Code for all of these classes is listed below:

```
/* SaneSlider.h - A slider with a working resize policy */

#ifndef _SaneSlider_
#define _SaneSlider_

#include <qslider.h>
#include <qsizepolicy.h>

#define MINSCALE 1
#define MAXSCALE 7
```

```
class SaneSlider : public QSlider
{
    Q_OBJECT
    public:
        SaneSlider(QWidget* parent, const char* name);
    private:
        QSize mySize;

};

#endif /* _SaneSlider_ */

/* SaneSlider.cpp - implementation for saneSlider class*/

#include "SaneSlider.h"
#include <stdio.h>


SaneSlider::SaneSlider(QWidget* parent,const char* name)
          : QSlider(QSlider::Horizontal, parent, name)
{
    setMinValue(MINSCALE);
    setMaxValue(MAXSCALE);
    setValue(MINSCALE);
    setPageStep(1);
    setLineStep(1);
    setTickmarks(QSlider::Below);
    mySize.setWidth(200);
    mySize.setHeight(25);
    setFixedSize(mySize);
}


/* RatingScale.h - A rating scale widget */

#ifndef _RatingScale_
#define _RatingScale_

#include <qwidget.h>
#include <qlabel.h>
#include <qfont.h>
#include "SaneSlider.h"

class RatingScale : public QWidget
{
    Q_OBJECT
    public:
        RatingScale(QWidget* parent, const char* name);
        void resetScale(const char* label);
        int getValue();
```

```
        const char* getLabel();
    private:
        QLabel* label;
        SaneSlider* slider;
        static QFont* labelFont;
};

#endif /* _RatingScale_ */

/* RatingScale.cpp - implementation for RatingScale class*/

#include "RatingScale.h"
#include <qlayout.h>
#include <qnamespace.h>

QFont* RatingScale::labelFont = new QFont(
    "Helvetica", 12, QFont::Normal
);

RatingScale::RatingScale(
    QWidget *parent,
    const char *name
) : QWidget (parent)
{
    QGridLayout* layout = new QGridLayout(this,1,2);
    layout->setColStretch(0,1);
    layout->setColStretch(1,0);

    label = new QLabel(this,"label");
    label->setFont(*labelFont);
    label->setAlignment(Qt::AlignLeft);
    layout->addWidget(label,0,0);

    slider = new SaneSlider(this,"slider");
    layout->setSpacing(20);
    layout->setMargin(5);
    layout->addWidget(slider,0,1);
}

void RatingScale::resetScale(const char* txt)
{
    label->setText(txt);
    slider->setValue(MINSCALE);
}

int RatingScale::getValue()
{
    return slider->value();
}

const char* RatingScale::getLabel()
```

```
{
    return label->text();
}


/* RateDisplay.h - The display window for ratings */

#ifndef _RateDisplay_
#define _RateDisplay_
#include <qwidget.h>
#include <qpushbutton.h>
#include <iostream.h>
#include "RatingScale.h"


class RateDisplay : public QWidget
{
    Q_OBJECT
    public:
        RateDisplay(QWidget* parent, const char* name,
                        int titles, int words);
        void displayMe(char** titles, char*** words);
        void writeResults(ostream &os);
    private slots:
        void hearButtonClick();
    signals:
        void buttonClicked();
    private:
        RatingScale** titleScales; //array of scales for titles
        RatingScale** wordScales; //array of scales for words
        RatingScale* integrateScale; //integration rating
        int numWords, numTitles;
        QPushButton* button;
        static QFont* titleFont;
};

#endif /* _RateDisplay_ */

/* RateDisplay.cpp - implementation for RateDisplay class */

#include "Experiment.h"
#include "RateDisplay.h"
#include <qlabel.h>
#include <qlayout.h>
#include <qnamespace.h>


QFont* RateDisplay::titleFont = new QFont(
    "Helvetica", 14, QFont::Bold
);
```

```
RateDisplay::RateDisplay(QWidget *parent, const char *name,
        int titles, int words)
        : QWidget (parent, name)
{
    numTitles = titles;
    numWords = words;

    //configure layout manager
    QGridLayout* layout = new QGridLayout(this,
        (words+1>>1)+titles+3+(Experiment::withMusic ? 2: 0), 2
    );

    //create RatingScales for titles and words
    titleScales = new RatingScale*[titles];
    for (int i=0; i<titles; i++)
        titleScales[i] = new RatingScale(this, name);

    wordScales = new RatingScale*[words];
    for (int i=0; i<words; i++)
        wordScales[i] = new RatingScale(this, name);
    if (Experiment::withMusic){
        integrateScale = new RatingScale(this, name);
    }

    //and labels & buttons
    QLabel* titleLabel = new QLabel(
        "How well do you feel these titles reflect or capture "
        "the photograph? \n(\"just about satisfactory\" <--> "
        "\"extremely well\")",
        this
    );
    titleLabel->setFont(*titleFont);
    titleLabel->setAlignment(Qt::AlignCenter|Qt::WordBreak);

    QLabel* wordsLabel = new QLabel(
        "How appropriate are these words to describe the "
        "photograph? \n(\"extremely inappropriate\" <--> "
        "\"extremely appropriate\")",
        this
    );
    wordsLabel->setFont(*titleFont);
    wordsLabel->setAlignment(Qt::AlignCenter|Qt::WordBreak);

    QLabel* integrateLabel = NULL;
    if(Experiment::withMusic){
        integrateLabel = new QLabel(
            "To what extent did the experience of listening to "
            "the music integrate with the experience of viewing "
            "the photograph?",
            this
```

```
        );
        integrateLabel->setFont(*titleFont);
        integrateLabel->setAlignment(Qt::AlignCenter|Qt::WordBreak);
    }

    button = new QPushButton
        ("Click here for next trial", this, "button");
    button->setFont(*titleFont);

    connect(
        button, SIGNAL(clicked()),
        this, SLOT(hearButtonClick())
    );

    //lay out components
    int z=0; //row
    layout->addMultiCellWidget(titleLabel,z,z,0,1); z++;
    for (int i=0; i<titles; i++){
        layout->addMultiCellWidget(titleScales[i],z,z,0,1);
        z++;
    }
    layout->addMultiCellWidget(wordsLabel,z,z,0,1); z++;
    for (int i=0; i<words; i++){
        layout->addWidget(wordScales[i],z,i&1);
        z+=(i&1);
    }
    z+=(words&1);
    if (Experiment::withMusic){
        layout->addMultiCellWidget(integrateLabel,z,z,0,1); z++;
        layout->addMultiCellWidget(integrateScale,z,z,0,1); z++;
    }
    layout->addMultiCellWidget(button,z,z,0,1);
}

/* SLOT */
void RateDisplay::hearButtonClick()
{
    button->setEnabled(false);
    this->hide();
    emit buttonClicked();
}

/* Resets each rating scale. Sets word scale labels
 * according to **words; ditton with **titles. Both passed
 * arrays MUST contained the right number of elements
 * otherwise behaviour is undefined, and probably not very
 * satisfactory...
 */
void RateDisplay::displayMe(char** titles, char*** words)
{
```

```
    for (int i=0; i<numTitles; i++)
        titleScales[i]->resetScale(titles[i]);
    for (int i=0; i<numWords; i++)
        wordScales[i]->resetScale(words[i][0]);
    if (Experiment::withMusic)
        integrateScale->resetScale("Integration");
    button->setEnabled(true);
    this->show();
}


/* Writes all values from the scales out to os */
void RateDisplay::writeResults(ostream &os)
{
    os << "Titles:\n";
    for (int i=0; i<numTitles; i++)
        os << titleScales[i]->getLabel()
           << "="
           << titleScales[i]->getValue()
           << "\n";
    os << "\nWords:\n";
    for (int i=0; i<numWords; i++)
        os << wordScales[i]->getLabel()
           << "="
           << wordScales[i]->getValue()
           << "\n";
    if (Experiment::withMusic)
        os << "\nIntegration:\n"
           << integrateScale->getLabel()
           << "="
           << integrateScale->getValue()
           << "\n";
    os << "\n";
}
```

All of the above code can be found on *Compact Disc One - Software*, along with a modified version of the Chapter Eight Makefile that will compile and link it. Having run "make" and "make setuid" as the *root* user, the program can be run using the command "exp3 <filename> <mus>" where <filename> is the name of a file to which the results will be output and <mus> is a value of "0" (no music) or "1" (with music).

In addition to the program described above, a simple stand-alone Java program was used to convert the session output files into a form suitable for analysis. The program outputs five different files: three for performing ANOVA calculations on title, word and integration data, along with two designed for correlation analyses. Pressure sensor data is downsampled to a single value using the same triangle function as that used in the equivalent program written for the Chapter Eight experiment (see above). The program code is listed in full below:

```
/**
 * Analysis Preparation tool for Experiment presented
 * in Chapter Nine. Input files and output filenames
 * are hardcoded below...
 */
package analysis;
import java.io.*;
import java.util.*;

public class DataPreparer
{
    //file locations etc
    public static final String datadir =
        "experiments/exp_3/data/";
    public static final String wordFile =
       "experiments/exp_3/analysis/words.txt";
    public static final String titleFile =
        "experiments/exp_3/analysis/titles.txt";
    public static final String outMainF = "/tmp/nMainAn.txt";
    public static final String outWordsF = "/tmp/nWordsAn.txt";
    public static final String outSensor1F = "/tmp/nSensor1Ans.txt";
    public static final String outSensor2F = "/tmp/nSensor2Ans.txt";
    public static final String outSensor3F = "/tmp/nSensor3Ans.txt";
    public static final String outPasActDiffF = "/tmp/nPasActDiff.txt";
    public static final String[] datafiles = {
        "M-hannah.txt","M-paul.txt","christian.txt","john.txt",
        "M-alice.txt","M-julian.txt","M-rachelb.txt","christophe.txt",
        "julie.txt","M-andreww.txt","M-katie.txt","M-tomc.txt",
        "gemma.txt","liz.txt","M-chrisj.txt","M-katieh.txt",
        "anthony.txt","helen.txt","naomi.txt","M-george.txt",
        "M-mark.txt","apt1002.txt","jenny.txt","rachelg.txt"
    };

    //basic parameters
    public static final int NUM_WORDS_PER_CAT = 6;
    public static final int NUM_TRIALS = 15;
    public static final int NUM_TITLES = 4;
    public static final int SILLY_NUMBER = -9999;
    public static boolean withMusic;

    //data
    public static List words;
    public static List pictures;
    public static Map titles;
```

```
/* Main program  */
public static void main(String[] args)
throws IOException
{
    //global initialisation
    titles = new HashMap();
    pictures = new ArrayList();
    loadTitleMapAndPictureList();
    words = new ArrayList();
    loadWordList();

    //build output files
    PrintWriter outMain = new PrintWriter(
        new FileWriter(outMainF));
    PrintWriter outWords = new PrintWriter(
        new FileWriter(outWordsF));
    PrintWriter outSensor1 = new PrintWriter(
        new FileWriter(outSensor1F));
    PrintWriter outSensor2 = new PrintWriter(
        new FileWriter(outSensor2F));
    PrintWriter outSensor3 = new PrintWriter(
        new FileWriter(outSensor3F));
    PrintWriter outPasActDiff = new PrintWriter(
        new FileWriter(outPasActDiffF));

    //headers
    String[] lbl = {"HN","HP","N"};
    String mainHdr="Group\t",wordsHdr="Group\t",s1Hdr="",s3Hdr="";
    for (int i=0;i<lbl.length; i++){ //HN/HP/N
        String pfx = lbl[i] + "_";
        mainHdr += pfx + "P_B" + "\t";
        mainHdr += pfx + "P_A" + "\t";
        mainHdr += pfx + "N_B" + "\t";
        mainHdr += pfx + "N_A" + "\t";
           wordsHdr += pfx + "N" + "\t";
        wordsHdr += pfx + "P" + "\t";
        s1Hdr += pfx + "I\t";
        s3Hdr += pfx + "S\t";
    }
    outMain.println(mainHdr);
    outWords.println(wordsHdr);
    outSensor1.println(s1Hdr);
    outSensor2.println("Int\tSens");
    outSensor3.println(s3Hdr);
    outPasActDiff.println(
       "Int\tSens\tPADiff\tWrdDiff\tPhoto\tMusic"
    );

    //now do it for each datafile
    for (int i=0; i<datafiles.length; i++){
```

```
//prepare
withMusic = datafiles[i].startsWith("M-");
TrialData[] ans = parseFile(datadir + datafiles[i]);
String mainD = (withMusic ? "Mus" : "NoMus") + "\t";
String wordsD = (withMusic ? "Mus" : "NoMus") + "\t";
String sens1D = "",sens3D = "";

//for each HP_HN_N category, sort out data
//5 TrialDatas gives us mean data for each cat
for (int j=0; j< lbl.length; j++){ //cat
    double[][] rScores = new double[2][2];
    double rNarWords=0;
    double rNonNarWords=0;
    double rInteg=0;
    double rSens=0;
    for (int k=0; k<5; k++){ //example
        //add up mean scores per title
        rScores[0][0] += ans[5*j+k].titleScore[0][0];
        rScores[0][1] += ans[5*j+k].titleScore[0][1];
        rScores[1][0] += ans[5*j+k].titleScore[1][0];
        rScores[1][1] += ans[5*j+k].titleScore[1][1];
        //tot up words
        double narWords=0,nonNarWords=0;
        for (int l=0; l<NUM_WORDS_PER_CAT;l++){
            narWords += ans[5*j+k].narWords[l];
             nonNarWords += ans[5*j+k].noNarWords[l];
        }
        rNarWords += narWords;
        rNonNarWords += nonNarWords;
        if (withMusic){
            rInteg += ans[5*j+k].integration;
            rSens += ans[5*j+k].meanPressure;
            outSensor2.println(
                ans[5*j+k].integration +
                "\t" +
                ans[5*j+k].meanPressure
            );
            outPasActDiff.println(
                ans[5*j+k].integration + "\t" +
                ans[5*j+k].meanPressure + "\t" +
                ( //pasActDifferences
                    ( ans[5*j+k].titleScore[1][0] + //act
                      ans[5*j+k].titleScore[1][1] )
                    -
                    ( ans[5*j+k].titleScore[0][0] + //pas
                      ans[5*j+k].titleScore[0][1] )
                 )/2.0 + "\t" +
                ( //narNonNarWrd differences
                    (narWords - nonNarWords)
                      / NUM_WORDS_PER_CAT
```

```
                        ) + "\t" + ans[5*j+k].imgName + "\t"
                          + ans[5*j+k].music
                    );
                }
            }

            //build outfile strings
            mainD += (rScores[0][0])/5.0 + "\t";
            mainD += (rScores[0][1])/5.0 + "\t";
            mainD += (rScores[1][0])/5.0 + "\t";
            mainD += (rScores[1][1])/5.0 + "\t";
            wordsD += rNarWords/(5.0*NUM_WORDS_PER_CAT) + "\t";
            wordsD += rNonNarWords/(5.0*NUM_WORDS_PER_CAT) + "\t";
            if (withMusic){
                sens1D += rInteg/5.0 + "\t";
                sens3D += rSens/5.0 + "\t";
            }
        }

        outMain.println(mainD);
        outWords.println(wordsD);
        if (withMusic){
            outSensor1.println(sens1D);
            outSensor3.println(sens3D);
        }
    }

    outMain.close();
    outWords.close();
    outSensor1.close();
    outSensor2.close();
    outSensor3.close();
    outPasActDiff.close();
}


/* Read a participant data file and fill up an
 * array of TrialDatas from it
 */
public static TrialData[] parseFile(String filename)
throws IOException
{
    TrialData[] answers = new TrialData[NUM_TRIALS];
    BufferedReader br = new BufferedReader(
        new FileReader(filename)
    );
    br.readLine();
    //for each trial
    for (int tr=0; tr<NUM_TRIALS;tr++){
       //create data structure
```

```
TrialData td = new TrialData();

br.readLine(); //trial num

//which picture are we dealing with?
StringTokenizer tok = new StringTokenizer(
    br.readLine(),
     "/", false
);
while(tok.countTokens()>1) tok.nextToken();
td.imgName = tok.nextToken();
if (td.imgName.startsWith("hp")) td.imgType = '+';
if (td.imgName.startsWith("hn")) td.imgType = '-';
if (td.imgName.startsWith("n")) td.imgType = '0';

br.readLine(); //blank
if (withMusic){
    //read music filename
    tok = new StringTokenizer(
        br.readLine(),
        "/",false
    );
    while(tok.countTokens()>1) tok.nextToken();
    td.music = tok.nextToken();
    br.readLine();
    tok = new StringTokenizer(
        br.readLine(),
        "\t",false
    );
    double[] sdata = new double[tok.countTokens()];
    for (int i=0; tok.hasMoreTokens(); i++){
        sdata[i] = Integer.parseInt(tok.nextToken());
    }
    td.meanPressure = DataFormatter.downSample(
        DataFormatter.cutTrailingZeros(sdata),1)
        [0];
}

br.readLine(); //titles
//read title scores
for (int i=0; i<NUM_TITLES; i++){
    tok = new StringTokenizer(
       br.readLine(),
       "=",false
    );
    String titleType = (String)titles.get(tok.nextToken());
    int score = Integer.parseInt(tok.nextToken());
    if (titleType.equals("PB")) td.titleScore[0][0] = score;
    if (titleType.equals("PA")) td.titleScore[0][1] = score;
    if (titleType.equals("NB")) td.titleScore[1][0] = score;
```

```
            if (titleType.equals("NA")) td.titleScore[1][1] = score;
        }

        br.readLine(); //blank line
        br.readLine(); //words
        //find word scores
        for (int i=0; i<(NUM_WORDS_PER_CAT*2);i++){
            tok = new StringTokenizer(
                br.readLine(),
                "=",false
            );
            //if its in the first half, its a noN
            int index = words.indexOf(tok.nextToken());
            if(index < NUM_WORDS_PER_CAT){
                td.narWords[index] = Integer.parseInt(tok.nextToken());
            } else {
                td.noNarWords[index - NUM_WORDS_PER_CAT] =
                    Integer.parseInt(tok.nextToken());
            }
        }

        br.readLine(); //blank line
        if (withMusic){
            br.readLine(); //integration
            tok = new StringTokenizer(
                br.readLine(),
                "=",false
            );
            tok.nextToken();
            td.integration = Integer.parseInt(tok.nextToken());
            br.readLine(); //blank line
        }

        //place structure in answers[]
        answers[pictures.indexOf(td.imgName)] = td;

    }
    return answers;
}


/* Load titles into hashmap, so its type can be found */
public static void loadTitleMapAndPictureList()
throws IOException
{
    BufferedReader br = new BufferedReader(
        new FileReader(titleFile)
    );
    for (int i=0; i<NUM_TRIALS; i++){
        StringTokenizer tok = new StringTokenizer(
```

```
                br.readLine(),",",false
            );

            //grab picture name
            StringTokenizer t = new StringTokenizer(
                   tok.nextToken(),
                "/", false
            );
            while(t.countTokens()>1) t.nextToken();
            pictures.add(t.nextToken());

            //[Passive/Narrative][Boring][Arousing]
            titles.put(tok.nextToken(),"PB");
            titles.put(tok.nextToken(),"PA");
            titles.put(tok.nextToken(),"NB");
            titles.put(tok.nextToken(),"NA");
        }
    }

    /* Load words into List, so their identity can be found */
    public static void loadWordList()
    throws IOException
    {
        BufferedReader br = new BufferedReader(
            new FileReader(wordFile)
        );
        for (int i=0; i<(NUM_WORDS_PER_CAT*2); i++){
            words.add(br.readLine());
        }
    }

}


/* This class encapsulates the data associated with a single
 * trial for a single participant
 */
class TrialData
{
    public TrialData(){
        titleScore = new int[2][2];
        noNarWords = new int[DataPreparer.NUM_WORDS_PER_CAT];
        narWords = new int[DataPreparer.NUM_WORDS_PER_CAT];
    }
    public String imgName;
    public char imgType; //+,-,0
    public String music;
    public int[][] titleScore; //[no nar/nar][plain/arouse]
    public int[] noNarWords;
    public int[] narWords;
```

```
    public int integration;
    public double meanPressure;

    public String toString()
    {
        String me =
            "ImageName=" + imgName + "\n" +
            "Music=" + music + "\n" +
            "Type=" + imgType + "\n" +
            "PB=" + titleScore[0][0] + "\n" +
            "PA=" + titleScore[0][1] + "\n" +
            "NB=" + titleScore[1][0] + "\n" +
            "NA=" + titleScore[1][1] + "\n" +
            "NarrWrds:";
        for (int i=0; i<narWords.length;i++) me += narWords[i] + ",";
        me += "\nNonNarrWrds:";
        for (int i=0; i<noNarWords.length;i++) me += noNarWords[i] + ",";
        me += "\n";
        me += "Integration=" + integration + "\n";
        me += "Pressure=" + meanPressure + "\n\n";
        return me;
    }
}
```

To use this program, modify the file locations to point to the appropriate place on the system on which the code is to be run; then compile and run using any JDK1.2-compliant compiler and virtual machine. The program takes no parameters as all are hard-coded.

For any further information about any of the code presented in this appendix, please contact the author of the thesis.

Appendix B

Stimulus Material

This appendix lists sources for stimulus material used in the experiments reported in Part Two of the thesis. Table B.1 lists the books in which the short stories used in the experiment in Chapter Seven were found; table B.2 lists the CDs from which the musical extracts used in that experiment were taken. Table B.3 lists the CDs from which the Chapter Eight musical stimuli were taken, and table B.4 those from which the Chapter Nine musical stimuli were taken. As the musical extracts listed in table B.3 and table B.4 were taken from quite specific points within the recordings, a copy of the recordings as they were used within the experiment are presented on *Compact Disc Two - Stimuli*. Track numbers on this CD correspond to the extract number listed in table B.3 and to the numbers in parentheses in table B.4.

Story	*Source*
The Window	Chandler, Robert (1986), *Just Imagine: Short Stories*,Rothesay: Writer's Rostrum.
Intaglio	Chandler, Robert (1986), *Just Imagine: Short Stories*, Rothesay: Writer's Rostrum.
The Old Man at the Bridge	Hemingway, Ernest (1986), *Short Stories*, Hughes (ed.), London: The Folio Society.
The Masque of the Red Death	Poe, Edgar Allan (1991), *The Gold-Bug and Other Tales*, Applebaum (ed.), New York: Dover Publications.
The Tell-Tale Heart	Poe, Edgar Allan (1991), *The Gold-Bug and Other Tales*, Abblebaum (ed.), New York: Dover Publications.
The Dress	Thomas, Dylan (1992), *Classic Welsh Short Stories*, Jones & Elis (eds.), Oxford: OUP
Cat in the Rain	Hemingway, Ernest (1986), *Short Stories*, Hughes (ed.), London: The Folio Society.
Marriage Lines	Lively, Penelope (1998), *Beyond the Blue Mountains*, London: Penguin.
The Dream Merchant	Lively, Penelope (1987), *Pack of Cards*, London: Penguin.
A Newspaper Story	Henry, O. (1992), *The Gift of the Magi and Other Short Stoires*, Weller (ed.), New York: Dover Publications.
The Voice of the City	Henry, O. (1992), *The Gift of the Magi and Other Short Stoires*, Weller (ed.), New York: Dover Publications.
Three Lambs	O'Flaherty, Liam (1985), *Classic Irish Short Stories*, O'Connor (ed.), Oxford: OUP.

Table B.1: Stories presented in Chapter Seven

Extract	*Source*
Beethoven, *Piano Concerto No. 4; Second Movement*	Maurizio Pollini (piano), Karl Bohm (cond.), Wiener Phlharmoniker. Deutsche Grammophon 439 483-2.
Brahms, *Piano Concerto No.1; Second Movement*	Julius Katchen (piano), Pierre Monteux (cond.), London Symphony Orchestra. Decca 440 612-2.
Rachmaninov, *Symphony No. 2; Third Movement*	Alexander Dmitriev (cond.), Symphony Orchestra of the St. Petersburg Philharmonia. Sony Classical SMK 57650.
Sibelius, *Symphony No.7; Opening*	Anthony Collins (cond.), London Symphony Orchestra. Decca LXT 2960.
Mahler, *Symphony No.3; First Movement*	Leif Segerstam (cond.), Danish National Radio Symphony Orchestra and Choir. Chandos CHAN 8970/1.
Schoenberg, *Verklarte Nacht; Opening*	Vladimir Ashkenazy (cond.), English Chamber Orchestra. Decca 410 111-2.
Mozart, *Piano Concerto No.23; Third Movement*	Daniel Barenboim (cond.), English Chamber Orchestra. EMI CZS 62825 2.
Shostakovich, *Festive Overture Op.96; Opening*	Neeme Jarvi (cond.), Scottish National Orchestra. Chandos CHAN 8587.
Rachmaninov, *Piano Concerto No.2; Second Movement*	Vladimir Ashkenazy (piano), Bernard Haitink (cond.), Concertgebouw Orchestra. Decca 421 590-2.
Handel, *Water Music Suite; Second Movement*	Trevor Pinnock (cond.), The English Concert. Polydor Archiv 423 149-2.
Bach, *Orchestral Suite No.3; Third Movement*	Philip Ledger (cond.), English Chamber Orchestra. Virgo VJ 91452-2.
Beethoven, *Symphony No.3; Third Movement*	Sir Colin Davis (cond.), Staatskapelle Dresden. Philips 446 824-2.

Table B.2: Musical extracts presented in Chapter Seven

Extract	*Source*
1	Track 2 of *The Film Music of Georges Auric.* Rumon Gamba (cond.), BBC Philharmonic. Chandos CHAN 9774.
2	Track 3 of *The Film Music of Georges Auric.* Rumon Gamba (cond.), BBC Philharmonic. Chandos CHAN 9774.
3	Track 7 of *The Film Music of Georges Auric.* Rumon Gamba (cond.), BBC Philharmonic. Chandos CHAN 9774.
4	Track 23 of *The Film Music of Georges Auric.* Rumon Gamba (cond.), BBC Philharmonic. Chandos CHAN 9774.
5	Track 27 of *The Film Music of Georges Auric.* Rumon Gamba (cond.), BBC Philharmonic. Chandos CHAN 9774.
6	Track 27 of *The Film Music of Georges Auric.* Rumon Gamba (cond.), BBC Philharmonic. Chandos CHAN 9774.
7	Track 1 of *Henry V Original Sountrack Recording.* Patrick Doyle (composer), Simon Rattle (cond.), City of Birmingham Symphony Orchestra. EMI CDC 49919 2.
8	Track 3 of *Henry V Original Sountrack Recording.* Patrick Doyle (composer), Simon Rattle (cond.), City of Birmingham Symphony Orchestra. EMI CDC 49919 2.
9	Track 6 of *Henry V Original Sountrack Recording.* Patrick Doyle (composer), Simon Rattle (cond.), City of Birmingham Symphony Orchestra. EMI CDC 49919 2.
10	Track 9 of *Henry V Original Sountrack Recording.* Patrick Doyle (composer), Simon Rattle (cond.), City of Birmingham Symphony Orchestra. EMI CDC 49919 2.
11	Track 14 of *Henry V Original Sountrack Recording.* Patrick Doyle (composer), Simon Rattle (cond.), City of Birmingham Symphony Orchestra. EMI CDC 49919 2.
12	Track 2 of *The Empire Strikes Back, Special Edition: The original motion picture soundtrack,* Disk 2, John Williams (composer & cond.), BMG Classics 09026-68747-2.
13	Track 3 of *The Empire Strikes Back, Special Edition: The original motion picture soundtrack,* Disk 2, John Williams (composer & cond.), BMG Classics 09026-68747-2.
14	Track 3 *of The Empire Strikes Back, Special Edition: The original motion picture soundtrack,* Disk 2, John Williams (composer & cond.), BMG Classics 09026-68747-2.
15	Track 5 *of The Empire Strikes Back, Special Edition: The original motion picture soundtrack,* Disk 2, John Williams (composer & cond.), BMG Classics 09026-68747-2.
16	Track 5 of *The Empire Strikes Back, Special Edition: The original motion picture soundtrack,* Disk 2, John Williams (composer & cond.), BMG Classics 09026-68747-2.
17	Track 7 of *The Empire Strikes Back, Special Edition: The original motion picture soundtrack,* Disk 2, John Williams (composer & cond.), BMG Classics 09026-68747-2.
18	Track 11 of *The Empire Strikes Back, Special Edition: The original motion picture soundtrack,* Disk 2, John Williams (composer & cond.), BMG Classics 09026-68747-2.
19	Track 12 of *Black Sunday Suite* from *Jurassic Park: The Classic John Williams.* William Motzing (cond.), City of Prague Philharmonic. Silva Screen Records FILMCD 147.
20	Track 12 of *Black Sunday Suite* from *Jurassic Park: The Classic John Williams.* William Motzing (cond.), City of Prague Philharmonic. Silva Screen Records FILMCD 147.

Table B.3: Musical extracts presented in Chapter Eight

Extract	*Source*
1 (12)	Track 2 of *The Empire Strikes Back, Special Edition: The original motion picture soundtrack*, Disk 2, John Williams (composer & cond.), BMG Classics 09026-68747-2. (Extract also used in Chapter Eight experiment).
2 (20)	Track 12 of *Black Sunday Suite* from *Jurassic Park: The Classic John Williams*. William Motzing (cond.), City of Prague Philharmonic. Silva Screen Records FILMCD 147. (Extract also used in Chapter Eight experiment).
3 (3)	Track 7 of *The Film Music of Georges Auric*. Rumon Gamba (cond.), BBC Philharmonic. Chandos CHAN 9774. (Extract also used in Chapter Eight experiment).
4 (4)	Track 23 of *The Film Music of Georges Auric*. Rumon Gamba (cond.), BBC Philharmonic. Chandos CHAN 9774. (Extract also used in Chapter Eight experiment).
5 (6)	Track 27 of *The Film Music of Georges Auric*. Rumon Gamba (cond.), BBC Philharmonic. Chandos CHAN 9774. (Extract also used in Chapter Eight experiment).
6 (9)	Track 6 of *Henry V Original Sountrack Recording*. Patrick Doyle (composer), Simon Rattle (cond.), City of Birmingham Symphony Orchestra. EMI CDC 49919 2. (Extract also used in Chapter Eight experiment).
7 (21)	Track 5 of *Henry V Original Sountrack Recording*. Patrick Doyle (composer), Simon Rattle (cond.), City of Birmingham Symphony Orchestra. EMI CDC 49919 2.
8 (22)	Track 13 of *Henry V Original Sountrack Recording*. Patrick Doyle (composer), Simon Rattle (cond.), City of Birmingham Symphony Orchestra. EMI CD
9 (23)	*Aria* (track 1) from *Goldberg Variationen*. J.S.Bach (composer), Tatiana Nikolaeva (piano). Bluebell ABCD 043.
10 (24)	Movement 1 (track 1) from *Concerto for Piano and Orchestra in D minor op.15*. Johannes Brahms (composer), Maurizio Pollini (piano), Claudio Abbado (cond.), Berliner Philharmoniker. Deutsche Grammophon 447 041-2.
11 (25)	*Dreamin' the Dream* (track 7) from *Time*. Fleetwood Mac. Warner Bros. 9362-45920-2.
12 (26)	*Man Out of Time* (track 5) from *Imperial Bedroom*. Elvis Costello and the Attractions. Demon Records DPAM 8.
13 (27)	*Ramifications for string orchestra* (track 5). György Ligeti (composer), Michael Cielen (cond.), Berlin Radio-Symphony Orchestra. Wergo WER 60162-50.
14 (28)	Variation X (track 10) from *Enigma Variations*. Sir Edward Elgar (composer). Simon Rattle (cond.), City of Birmingham Symphony Orchestra. EMI CDC 5 55001 2.
15 (29)	*Warning* (track 1). The Levellers. WOL CD1034.

Table B.4: Musical extracts presented in Chapter Nine

Bibliography

Abbate, C. and R. Parker (1989). *Analyzing opera: Verdi and Wagner*. Berkeley: University of California Press.

Ackerman, D. (1990). *A Natural History of the Senses*. London: Chapmans.

Adolphs, R. and D. Tranel (1999). Intact recognition of emotional prosody following amygdala damage. *Neuropsychologia 37*(11), 1285–1292.

Agawu, V. K. (1991). *Playing with Signs*. Princeton, NJ: Princeton University Press.

Agawu, V. K. (1999). The challenge of semiotics. In N. Cook and M. Everist (Eds.), *Rethinking Music*. Oxford: Oxford University Press.

Akstein, D. (1973). Terpsichoretrancetherapy: a new hypnopsychotherapeutic method. *International Journal of Clinical and Experimental Hypnosis 21*(3), 131–143.

Albersnagel, F. A. (1988). Velten and musical mood induction procedures: a comparison with accessibility of thought associations. *Behaviour Research and Therapy 26*(1), 79–96.

Anderson, J. R. and G. H. Bower (1973). *Human Associative Memory*. The Experimental Psychology Series. New York: John Wiley and Sons.

Arom, S. (1991). *African Polyphony and Polyrhythm: musical structure and methodology*. Cambridge: Cambridge University Press. Translated by Martin Thom, Barbara Tuckett and Raymond Boyd.

Aures, W. (1985). Berechnungsverfahren für den sensorischen wohlklang beliebiger schallsignale [procedure for estimating the pleasantness of the sensory quality of a resounding signal]. *Acustica 59*, 130–141.

Balch, W. R. and B. S. Lewis (1996). Music-dependent memory: the roles of tempo change and mood mediation. *Journal of Experimental Psychology: Learning, Memory, and Cognition 22*(6), 1354–1363.

Balkwill, L. L. and W. F. Thompson (1999). A cross-cultural investigation of the perception of emotion in music: psychophysical and cultural cues. *Music Perception 17*(1), 43–64.

Barker, R. (1991). *Neuroscience: an illustrated guide*. New York: Ellis Horwood.

Baumgartner, H. (1992). Remembrance of things past: music, autobiographical memory, and emotion. *Advances in Consumer Research 19*, 613–620.

Berlyne, D. (1971). *Aesthetics and Psychobiology*. New York: Appleton-Century-Crofts.

Bharucha, J. and P. Todd (1989). Modeling the perception of tonal structure with neural nets. *Computer Music Journal 13*(4), 44–53.

Bhatti, S. and A. Gregory (2000, August). Cross-cultural study of affective responses to Qawwali. In C. Woods, G. Luck, R. Brochard, F. Seddon, and J. Sloboda (Eds.), *Sixth International Conference on Music Perception and Cognition*.

Blacking, J. (1995). *Music, Culture and Experience: selected papers of John Blacking*. Chicago studies in ethnomusicology. Chicago: University of Chicago Press.

Boethius, A. M. S. (1989). *Fundamentals of Music*. New Haven: Yale University Press.

Bohlman, P. V. (2000). Ethnomusicology and music sociology. In D. Greer (Ed.), *Musicology and Sister Disciplines: Past, Present and Future*. Oxford: Oxford University Press.

Boltz, M., M. Schulkind, and S. Kantra (1991). Effects of background music on the remembering of filmed events. *Memory and Cognition 19*(6), 593–606.

Bordwell, D. (1985). *Narration in the Fiction Film*. London: Methuen.

Boris, H. N. (1993). *Passions of the Mind: unheard melodies: the third principle of mental functioning*. New York: New York University Press.

Bower, G. H. (1981). Mood and memory. *Americal Psychologist 36*, 129–148.

Bradley, M. and P. J. Lang (1999). Affective norms for English words (ANEW): instruction manual and affective ratings. Technical report, The Center for Research in Psychophysiology, University of Florida, Gainsville, Fl. Technical Report C-1.

Branigan, E. (1992). *Narrative Comprehension and Film*. London: Routledge.

Bregman, A. S. (1981). Asking the "what for" question in auditory perception. In M. Kubovy and J. Pomerantz (Eds.), *Perceptual Organization*. Hillsdale, New Jersey: Lawrence Erblaum.

Bregman, A. S. (1990). *Auditory Scene Analysis: the perceptual organization of sound*. Cambridge, MA: MIT Press.

Bruner, G. C. (1990). Music, mood, and marketing. *Journal of Marketing 54*(4), 94–104.

Bruner, J. (1986). *Actual Minds, Possible Worlds*. Cambridge,MA: Harvard University Press.

Buck, R. (1980). Nonverbal behavior and the theory of emotion: the facial feedback hypothesis. *Journal of Personality and Social Psychology 5*, 811–824.

Bullerjahn, C. and M. Gueldenring (1994). An empirical investigation of effects of film music using qualitiative content analysis. *Psychomusicology 13*(1-2), 99–118.

Bunt, L. (1994). *Music Therapy: An Art Beyond Words*. London: Routledge.

Burt, J. L., D. S. Bartolome, D. W. Burdette, and J. R. Comstock (1995). A psychophysiological evaluation of the perceived urgency of auditory warning signals. *Ergonomics 38*(11), 2327–2340.

Butler, D. (1989). Describing the perception of tonality in music: a critique of the tonal hierarchy theory and a proposal for a theory of intervallic rivalry. *Music Perception 6*(3), 219–241.

Carroll, N. (1997). Art, narrative, and emotion. In M. Hjort and S. Laver (Eds.), *Emotion and the Arts*. New York: Oxford University Press.

Center for the Study of Emotion and Attention [CSEA-NIMH] (1999). *International Affective Picture System: Digitized photographs*. Gainsville, Fl: The Center for Research in Psychophysiology, University of Florida.

Citron, S. (1985). *Song Writing: A Complete Guide to the Craft.* London: Hodder and Stoughton.

Claperede, E. (1951). Recognition and "me-ness". In D. Rapaport (Ed.), *Organization and Pathology of Thought.* New York: Columbia University Press.

Clarke, E. F. (1993). Generativity, mimesis and the human body in music performance. *Contemporary Music Review* 9(1-2), 207–219.

Clore, G. L. (1994). Why emotions require cognition. In P. Ekman and R. J. Davidson (Eds.), *The Nature of Emotions.* New York: Oxford University Press.

Clynes, M. (1977). *Sentics: the Touch of the Emotions.* New York: Anchor.

Clynes, M. and N. Nettheim (1982). The living quality of music: Neurobiologic patterns of communicating feeling. In M. Clynes (Ed.), *Music, Mind and Brain.* New York: Plenum Press.

Clynes, M. and J. Walker (1982). Neurobiologic functions of rhythm, time and pulse in music. In M. Clynes (Ed.), *Music, Mind and Brain.* New York: Plenum Press.

Cohen, A. J. (1999). The functions of music in multimedia: a cognitive approach. In S. W. Yi (Ed.), *Music, Mind, and Science.* Seoul: Seoul National University Press.

Cook, N. (1998). *Analysing Musical Multimedia.* Oxford: Clarendon Press.

Cooke, D. (1959). *The Language of Music.* Oxford: Oxford University Press.

Cross, I. R. (1995). The analysis and cognition of melodic complexity by Eugene Narmour. *Music Perception* 12(2), 486–509.

Crowder, R. G., M. L. Serafine, and B. Repp (1990). Physical interaction and association by contiguity in memory for the words and melodies of songs. *Memory and Cognition* 18(5), 469–476.

Damasio, A. R. (1994). *Descarte's Error: Emotion, Reason and the Human Brain.* New York: G.P. Putnam's Sons.

Dannenbring, G. (1976). Perceived auditory continuity with alternately rising and falling frequency transitions. *Canadian Journal of Psychology 30,* 99–114.

Darwin, C. (1872). *The Expression of the Emotions in Man and Animals*. London: Murray.

Darwin, C. and R. Carlyon (1995). Auditory grouping. In B. C. Moore (Ed.), *Hearing*. San Diego: Academic Press.

Davidson, J. W. (1991). *The perception of expressive movement in music performance*. Ph. D. thesis, City University, London.

Davidson, R. J. (1994). How are emotions distinguished from moods and other affective constructs? In P. Ekman and R. J. Davidson (Eds.), *The Nature of Emotions*. New York: Oxford University Press.

Davies, S. (1997). Contra the hypothetical persona in music. In M. Hjort and S. Laver (Eds.), *Emotion and the Arts*. New York: Oxford University Press.

Deliège, I., M. Mélen, D. Stammers, and I. Cross (1996). Musical schemata in real-time listening to a piece of music. *Music Perception 14*, 117–160.

Dowling, W. J. and D. L. Harwood (1986). *Music Cognition*. San Diego: Academic Press.

Eich, E. (1995). Searching for mood dependent memory. *Psychological Science 6*(2), 67–75.

Eich, E. and L. Ryan (2000). Mood dependence and implicit memory. In E. Tulving (Ed.), *Memory, consciousness, and the brain: The Tallinn Conference*, Philadelphia. Psychology Press.

Eisenberg, N. and J. Strayer (1986). *Empathy and its Development*. Cambridge: Cambridge University Press.

Ekman, P. (1992). Facial expressions of emotion: new findings, new questions. *Psychological Science 3*(1), 34–38.

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

Ekman, P. (1994a). Moods, emotions, and traits. In P. Ekman and R. J. Davidson (Eds.), *The Nature of Emotions*. New York: Oxford University Press.

Ekman, P. (1994b). Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique. *Psychological Bulletin 115*, 268–287.

Ellsworth, P. C. (1994). Some reasons to expect universal antecedents of emotion. In P. Ekman and R. J. Davidson (Eds.), *The Nature of Emotions*. New York: Oxford University Press.

Eysenck, M. W. and M. T. Keane (1995). *Cognitive Psychology: a student's handbook*. East Sussex: Psychology Press.

Feld, S. (1984). Communication, music, and speech about music. *Yearbook for Traditional Music 16*, 1–18.

Feld, S. (1990). *Sound and Sentiment: birds, weeping, poetics, and song in Kaluli expression*. Philadelphia: University of Pennsylvania Press.

Fenlon, I. and P. N. Miller (1992). *The Song of the Soul: Understanding Poppea*. London: Royal Musical Association.

Fernald, A. (1992). Meaningful melodies in mothers' speech to infants. In H. Papousek, U. Jurgens, and M. Papousek (Eds.), *Nonverbal Vocal Communication: Comparative and Developmental Aspects*. Cambridge: Cambridge University Press.

Finegan, J. E. and C. Seligman (1995). In defense of the velten mood induction procedure. *Canadian Journal of Behavioral Science 27*. Online Journal.

Finnäs, L. (1989). A comparison between young people's privately and publicly expressed musical preferences. *Psychology of Music 17*(2), 132–45.

Fodor, J. A. (1983). *Modularity of Mind*. Cambridge, MA: MIT Press.

Fodor, J. A. (1998). *Concepts: where cognitive science went wrong*. New York: Oxford University Press.

Fontana, A. E. and J. A. Loschi (1979). Combined use of music with sound of heart beats and respiratory rhythm in psychotherapy. *Acta Psiquiatrica y Psicologica de Americana Latina 25*(1), 34–41.

Freed, D. (1990). Auditory correlates of perceived mallet hardness for a set of recorded percussive sound events. *Journal of the Acoustical Society of America 87*, 311–322.

Frey, W. H. (1985). *Crying: the mystery of tears*. Minneapolis: Winston Press Inc.

Frick, R. (1985). Communicating emotion: the role of prosodic features. *Psychological Bulletin 97*, 412–429.

Frijda, N. H. (1986). *The Emotions*. Cambridge: Cambridge University Press.

Gabrielsson, A. (1973). Adjective ratings and dimension analyses of auditory rhythm patterns. *Scandinavian Journal of Psychology 14*(4), 244–260.

Gabrielsson, A. (1993). The complexities of rhythm. In T. J. Tighe and W. J. Dowling (Eds.), *Psychology and Music: the understanding of melody and rhythm*. Hillsdale, New Jersey: Lawrence Erblaum.

Gabrielsson, A. (1999). The performance of music. In D. Deutsch (Ed.), *The Psychology of Music* (Second ed.). San Diego: Academic Press.

Gabrielsson, A. and P. N. Juslin (1996). Emotional expression in music performance: Between the performer's intention and the listener's experience. *Psychology of Music 24*, 68–91.

Gabrielsson, A. and S. Lindström (2000). Strong experiences of and with music. In D. Greer (Ed.), *Musicology and Sister Disciplines: Past, Present and Future*. Oxford: Oxford University Press.

Gallup Jr., G. G. (1998). Self-awareness and the evolution of social intelligence. *Behavioral Processes 42*(2-3), 239–247.

Gavin, W. W. and G. Mandler (1987). Play it again, sam: on liking music. *Cognition and Emotion 1*(3), 259–282.

Gerra, G., A. Zaimovic, D. Franchini, M. Palladino, G. Guicastro, N. Reali, D. Maestri, R. Caccavari, R. Delsignore, and F. Brambilla (1998). Neuroendocrine responses of healthy volunteers to "techno-music": relationships with personality traits and emotional state. *International Journal of Psychophysiology 28*(1), 99–111.

Gerrig, R. J. (1989). Suspense in the absence of uncertainty. *Journal of Memory and Language 28*, 633–648.

Gerrig, R. J. (1993). *Experiencing Narrative Worlds: on the psychological activities of reading*. New Haven: Yale University Press.

Gfeller, K., E. Asmus, and M. Eckert (1991). An investigation of emotional response to music and text. *Psychology of Music 19*, 128–141.

Giomo, C. J. (1993). An experimental study of children's sensitivity to mood in music. *Psychology of Music 21*, 141–162.

Gratier, M. (2000). Expressions of belonging: the effect of acculturation on the rhythm and harmony of mother-infant vocal interaction. *Musicae Scientiae Special Issue, 1999/2000*, 93–122.

Gregory, A. H. (1997). The roles of music in society: the ethnomusicological perspective. In D. J. Hargreaves and A. C. North (Eds.), *The Social Psychology of Music*. Oxford: Oxford University Press.

Greiser, D. and P. Kuhl (1988). Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental Psychology 24*, 14–20.

Grey, J. (1977). Multidimensional perceptual scaling of musical timbres. *Journal of the Acoustical Society of America 61*(5), 1270–1277.

Grossenbacher, P. (1996). Consciousness and evolution in neocortex. In P. Mellars and K. Gibson (Eds.), *Modelling the Early Human Mind*. Cambridge: McDonald Institute Monographs.

Gundlach, R. (1932). A quantative analysis of Indian Music. *American Journal of Psychology 44*, 1–29.

Hajda, J. M., R. A. Kendall, E. C. Carterette, and M. L. Harshberger (1997). Methodological issues in timbre research. In I. Deliège and J. A. Sloboda (Eds.), *Perception and Cognition of Music*. Hove,East Sussex: Psychology Press.

Hakanen, E. A. (1995). Emotional use of music by African American adolescents. *Howard Journal of Communications 5*(3), 214–222.

Halpern, D., R. Blake, and J. Hillenbrand (1986). Psychoacoustics of a chilling sound. *Perception and Psychophysics 39*, 77–80.

Handel, S. (1995). Timbre perception and auditory object identification. In B. C. Moore (Ed.), *Hearing*. San Diego: Academic Press.

Hanslick, E. (1891). *On the Beautiful in Music: A contribution to the revival of musical aesthetics*. London: Novello. Translated from the 7th edition by Gustav Cohen.

Hevner, K. (1936). Experimental studies of the elements of expression in music. *American Journal of Psychology 48*, 248–268.

Holbrook, M. B. and R. M. Schindler (1989). Some exploratory findings on the development of musical tastes. *Journal of Consumer Research 16*(1), 119–124.

Huron, D. (1990a). Crescendo diminuendo asymmetries in beethoven piano sonatas. *Music Perception 7*(4), 395–402.

Huron, D. (1990b). Increment decrement asymmetries in polyphonic sonorities. *Music Perception 7*(4), 385–393.

Huron, D. (1991). The ramp archetype: A score-based study of musical dynamics in 14 piano composers. *Psychology of Music 19*(1), 33–45.

Huron, D. (1992). The ramp archetype and the maintenance of passive auditory attention. *Music Perception 10*(1), 83–92.

Isen, A. (1987). Positive affect, cognitive processes and social behavior. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology*, Volume 20. New York: Academic Press.

Iwanaga, M. and M. Tsukamoto (1997). Effects of excitative and sedative music on subjective and physiological relaxation. *Perception and Motor Skills 85*(1), 287–296.

Jackendoff, R. (1992). Musical processing and musical affect. In M. Jones and S. Holleran (Eds.), *Cognitive Bases of Musical Communication*. Washington: American Psychological Association.

Jarrell, T., C. Gentile, L. Romanski, P. McCabe, and N. Schneiderman (1987). Involvement of cortical and thalamic auditory regions in retention of differential bradycardia conditioning to acoustic conditioned stimuli in rabbits. *Brain Research 412*, 285–294.

Jaskowski, P., K. Rybarczyk, and F. Jaroszyk (1994). The relationship between latency of auditory evoked potentials, simple reaction time, and stimulus intensity. *Psychological Research 56*(2), 59–65.

Jaskowski, P., K. Rybarczyk, F. Jaroszyk, and D. Lemanski (1995). The effect of stimulus intensity on force output in simple reaction time task in humans. *Acta Neurobiologiae Experimentalis 55*(1), 57–64.

José, P. and W. Brewer (1984). Development of story-liking: Character identification, suspense, and outcome resolution. *Developmental Psychology 20*, 911–924.

Juslin, P. N. (1997). Emotional communication in music performance. *Music Perception 14*(4), 383–418.

Kappas, A., U. Hess, and K. R. Scherer (1991). Voice and emotion. In R. S. Feldman and B. Rime (Eds.), *Fundamentals of Nonverbal Behavior: Studies in Emotion and Social Interaction*. Cambridge: Cambridge University Press.

Kastner, M. P. and R. G. Crowder (1990). Perception of the major/minor distinction: iv. emotional connotations in young children. *Music Perception 8*(2), 189–201.

Keats, L. (1995). Doug: The rhythm in his world. *Canadian Journal of Music Therapy 3*(1), 53–69.

Kemp, A. E. (1997). Individual differences in musical behaviour. In D. J. Hargreaves and A. C. North (Eds.), *The Social Psychology of Music*. Oxford: Oxford University Press.

Kennedy, M. (1995). *Richard Strauss*. The Master Musicians. Oxford: Oxford University Press.

Kivy, P. (1980). *The Corded Shell: reflections on musical expression*. Princeton: Princeton University Press.

Kivy, P. (1984). *Sound and Semblance: reflections on musical representation*. Princeton, New Jersey: Princeton University Press.

Kivy, P. (1989). *Sound Sentiment: an essay on the musical emotions*. Philadelphia: Temple University Press.

Koffka, K. (1935). *Principles of Gestalt Psychology*. New York: Harcourt Brace.

Konečni, V. (1975). The mediation of aggressive behavior: arousal level vs anger and cognitive labeling. *Journal of Personality and Social Psychology 32*, 706–712.

Konečni, V., J. Crozier, and A. Doob (1976). Anger and expression of aggression: effects on aesthetic preference. *Scientific Aesthetics/Sciences de l'art 1*, 47–55.

Konečni, V. J. (1982). Social interaction and musical preference. In D. Deutsch (Ed.), *Psychology of Music* (First ed.). New York: Academic Press.

Kramer, L. (1984). *The Nineteenth Century and After*. Berkeley: University of California Press.

Kratus, J. (1993). A developmental study of children's interpretation of emotion in music. *Psychology of Music 21*(1), 3–19.

Krumhansl, C. L. (1996). A perceptual analysis of Mozart's Piano Sonata K.282: Segmentation, tension, and musical ideas. *Music Perception 13*(3), 401–432.

Krumhansl, C. L. (1997a). Can dance reflect the structural and expressive qualities of music? a perceptual experiment on balanchine's choreography of mozart's divertimento no.15. *Musicae Scientiae 1*(1), 63–85.

Krumhansl, C. L. (1997b). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology 51*, 336–352.

Krumhansl, C. L. (1998). Topic in music: an empirical study of memorability, openness and emotion in Mozart's String Quintet in C major and Beethoven's String Quartet in A minor. *Music Perception 16*(1), 119–134.

Labov, W. (1972). *Language in the Inner City: Studies in the Black English Vernacular*. Philadelphia: University of Pennsylvania Press.

Lakoff, G. (1987). *Women, Fire and Dangerous Things: What categories reveal about the mind*. Chicago: University of Chicago Press.

Lakoff, G. and M. Johnson (1980). *Metaphors We Live By*. Chicago: University of Chicago Press.

Lang, P. J., M. Bradley, and B. Cuthbert (1999). International affective picture system (IAPS): Instruction manual and affective ratings. Technical report, The Center for Research in Psychophysiology, University of Florida, Gainsville, Fl. Technical Report A-4.

Langer, S. K. (1951). *Philosophy in a New Key: A study in the symbolism of reason, rite and art.* London: Oxford University Press.

Langer, S. K. (1953). *Feeling and Form.* London: Routledge and Kegan Paul.

Lavy, M. M. (1999). Mupsych: computer-based experiments for music psychology. *Psychology Software News 9*(2), 41–42.

Lazarus, R. S. (1991). *Emotion and Adaptation.* New York: Oxford University Press.

Lazarus, R. S. (1994). The past and the present in emotion. In P. Ekman and R. J. Davidson (Eds.), *The Nature of Emotions.* New York: Oxford University Press.

LeDoux, J. (1998). *The Emotional Brain: The mysterious underpinnings of emotional life.* London: Weidenfeld and Nicolson.

LeDoux, J. and M. Rogan (1999). Emotion and the animal brain. In R. A. Wilson and F. Keil (Eds.), *The MIT Encyclopedia of the Cognitive Sciences.* Cambridge,MA: MIT Press.

LeDoux, J. E., C. Farb, and D. Ruggiero (1990). Topographic organization of neurons in the acoustic thalamus that project to the amygdala. *Journal of Neuroscience 10*, 1043–1054.

LeDoux, J. E., A. Sakaguchi, J. Iwata, and D. Reis (1984). Interruption of projections from the medial geniculate body to an archi-neostriatal field disrupts the classical conditioning of emotional responses to acoustic stimuli in the rat. *Neuroscience 17*, 615–627.

Lerdahl, F. and R. Jackendoff (1983). *A Generative Theory of Tonal Music.* Cambridge: MIT Press.

Levinson, J. (1997). Emotion in response to art: A survey of the terrain. In M. Hjort and S. Laver (Eds.), *Emotion and the Arts.* New York: Oxford University Press.

Lewis, J. M. (1994). Crowd crushes at two rock concerts: a value-added analysis. In J. S. Epstein (Ed.), *Adolescents and their Music: if it's too loud, you're too old.* New York: Garland Publishing.

Li, X., R. Logan, and R. Pastore (1991). Perception of acoustic source characteristics:walking sounds. *Journal of the Acoustical Society of America 90*(6), 3036–3049.

Lipps, T. (1913). *Leitfaden der Psychologie [Golden threads of Psychology]*. Leipzig: Engelmann.

Lowis, M. J. (1998). Music and peak experiences. *Mankind Quarterly 39*(2), 203–224.

MacCurdy, J. (1925). *The Psychology of Emotion*. New York: Harcourt, Brace and Co.

Madsen, C. K. and W. Fredrickson (1993). The experience of musical tension: a replication of nielsen's research using the continuous response digital interface. *Journal of Music Therapy 30*, 46–63.

Magrini, T. (2000). From music-makers to virtual singers: new musics and puzzled scholars. In D. Greer (Ed.), *Musicology and Sister Disciplines: Past, Present and Future*. Oxford: Oxford University Press.

Maher, T. (1980). A rigorous test of the proposition that musical intervals have different psychological effects. *American Journal of Psychology 93*(2), 309–327.

Mahler, A. (1968). *Gustav Mahler: memories and letters*. London: Murray. translated by Basil Creighton.

Malloch, S. N. (2000). Mothers and infants and communicative musicality. *Musicae Scientiae Special Issue, 1999/2000*, 59–74.

Mandler, G. (1984). *Mind and Body*. New York: Norton.

Mason, B. (1988). *Spence and Lila*. New York: Harper and Row.

Mathews, M. V. (1979). Perception of crescendos as a function of duration. *Journal of the Acoustic Society of America 65*(Supplement 1), S123.

Mauss, F. (1988). Music as drama. *Music Theory Spectrum 10*, 56–73.

McAdams, S. (1993). Recognition of sound sources and events. In S. McAdams and E. Bigand (Eds.), *Thinking in Sound*. Oxford: Clarendon Press.

McAdams, S. and J. Bertoncini (1997). Organization and discrimination of repeating sound sequences by newborn infants. *Journal of the Acoustical Society of America 102*(5,Pt1), 2945–2953.

McClary, S. (1991). *Feminine Endings: Music, gender, and sexuality*. Minnesota: University of Minnesota Press.

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

Meyer, L. B. (1961). On rehearing music. *Journal of the American Musicological Society 14*(2), 257–267.

Moore, B. C. (1989). *An Introduction to the Psychology of Hearing* (Third ed.). San Diego: Academic Press.

Morris, J., S. Scott, and R. Dolan (1999). Saying it with feeling: neural responses to emotional vocalizations. *Neuropsychologia 37*(10), 1155–1163.

Morton, E. (1977). On the occurrence and significance of motivational-structural rules in some bird and mammal sounds. *American Naturalist 111*, 855–69.

Narmour, E. (1990). *The Analysis and Cognition of Basic Melodic Structures: the Implication-realization model*. Chicago: Chicago University Press.

Narmour, E. (1992). *The Analysis and Cognition of Melodic Complexity*. Chicago: Chicago University Press.

Newcomb, A. (1983). Those images that yet fresh images beget. *Journal of Musicology 2*, 227–245.

Nielsen, F. (1983). *Oplevelse af musikalsk spænding [The experience of musical tension]*. Copenhagen: Akademisk Forlag.

Nielzen, S. and O. Olsson (1993). The influence of duration on verbal-attribute ratings of complex protomusical sounds. *Music Perception 11*(1), 73–86.

Nisbitt, R. and T. D. Wilson (1977). Telling more than we know: verbal reports on mental processes. *Psychological Review 84*(3), 231–259.

Nissen, M. J. (1977). Stimulus intensity and information processing. *Perception and Psychophysics 22*, 338–352.

North, A. C. and D. J. Hargreaves (2000). Music preference in everyday life: arousal- or typicality-based goals? Paper presented at "The effects of music", a conference of the Society for Research in Psychology of Music and Music Education, 8-9 April 2000.

North, A. C., D. J. Hargreaves, and J. McKendrick (1999). The influence of in-store music on wine selections. *Journal of Applied Psychology 84*(2), 271–276.

Noy, P. (1993). How music conveys emotion. In S. Feder, R. L. Karmel, and G. H. Pollock (Eds.), *Psychoanalytic Explorations in Music*. Madison, Connecticut: International Universities Press.

Oatley, K. (1992). *Best Laid Schemes: the Psychology of Emotions*. Cambridge: Cambridge University Press.

Ortony, A., G. Clore, and A. Collins (1988). *The Cognitive Structure of Emotions*. New York: Cambridge University Press.

Osborne, J. (1981). The mapping of thoughts, emotions, sensations and images as responses to music. *Journal of Mental Imagery 5*, 133–136.

Panksepp, J. (1997). The affective cerebral consequence of music: Happy vs sad effects on the EEG and clinical implications. *International Journal of Arts Medicine 5*(1), 18–27.

Papaeliou, C. and C. Trevarthan (1998). The infancy of music. *Musical Praxis 1*(2), 19–33.

Papoušek, H. (1996). Musicality in infant resarch. In I. Deliège and A. Sloboda, John (Eds.), *Musical Beginnings*. Oxford: Oxford University Press.

Pavlov, I. (1927). *Conditioned Reflexes*. London: Oxford University Press.

Peirce, C. S. (1931-1935). *Collected Papers*. Harvard, MA: Harvard University Press.

Peretz, I., L. Gagnon, and B. Bouchard (1998). Music and emotion: perceptual determinants, immediacy, and isolation after brain damage. *Cognition 68*, 111–141.

Pierce, J. R. (1983). *The Science of Musical Sound*. New York: W.H. Freeman.

Pillemer, D. B. (1993). Remembering personal circumstances: a functional analysis. In E. Winograd and U. Neisser (Eds.), *Affect and accuracy in recall: studies of " flashbulb " memories*. New York: Cambridge University Press.

Pinker, S. (1997). *How the Mind Works*. New York: W.W. Norton.

Pinkney, L. (1999). A comparison of the snoezelen environment and a music relaxation group on the mood and behaviour of patients with senile dementia. *British Journal of Occupational Therapy 60*(5), 209–212.

Piston, W. (1955). *Orchestration*. New York: W.W. Norton.

Plomp, R. (1976). *Aspects of tone sensation*. New York: Academic Press.

Pollard, H. and E. Janson (1982). A tristimulus method for the specification of musical timbre. *Acustica 51*, 162–171.

Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends in Cognitive Sciences 1*(2), 56–61.

Proust, M. (1934). *Remembrance of things past*, Volume 1. New York: Random House. Translated by C.K. Scott Moncrieff.

Putnam, L. and W. Roth (1990). Effects of stimulus repetition, duration, and rise time on startle blink and automatically elicited p300. *Psychophysiology 27*(3), 275–297.

Quittner, A. and R. Glueckhauf (1983). The facilitative effects of music on visual imagery: a multiple measures approach. *Journal of Mental Imagery 7*, 105–120.

Repp, B. H. (1997). The aesthetic quality of a quantitatively average music performance: two preliminary experiments. *Music Perception 14*(4), 419–444.

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

Rigg, M. G. (1964). The mood effects of music: a comparison of data from four investigators. *Journal of Psychology 58*, 427–438.

Rimè, B. and L. Schiaratura (1991). Gesture and speech. In R. S. Feldman and B. Rime (Eds.), *Fundamentals of Nonverbal Behavior: Studies in Emotion and Social Interaction*. Cambridge: Cambridge University Press.

Robb, L. (2000). Emotional musicality in mother-infant vocal affect, and an acoustic study of postnatal deptression. *Musicae Scientiae Special Issue, 1999/2000*, 123–154.

Robertson, E. (1934). The emotional element in listening to music. *Australian Journal of Psychology 12*, 199–212.

Ross, M. (1987). Popular Music. In J. Borwick (Ed.), *Sound Recording Practice*. Oxford: Oxford University Press.

Rouget, G. (1980). *Music and Trance: a theory of the relations between Music and Possession*. Chicago: University of Chigaco Press. Originally published as *la musique et la transe: esquisse d'une théorie générale des relations de la musique et de la passion*, Editions Gallimard, 1980.

Russell, J. A. (1979). Affective space is bipolar. *Journal of Personality and Social Psychology 37*(3), 345–356.

Russell, P. A. (1997). Musical tastes and society. In D. J. Hargreaves and A. C. North (Eds.), *The Social Psychology of Music*. Oxford: Oxford University Press.

Sakamoto, H., F. Hayashi, M. Tsujikawa, and S. Sugiuri (1997). Psycho-physiological responses by listening to some sounds from our daily life. *Journal of Sound and Vibration 205*(4), 499–503.

Schachter, S. and J. Singer (1962). Cognitive, social and physiological determinants of emotional state. *Psychological Review 69*, 379–399.

Schellenberg, E. (1996). Expectancy in melody: tests of the implication-realization model. *Cognition 58*, 75–125.

Schellenberg, E. (1997). Simplifying the implication-realization model of musical expectancy. *Music Perception 14*(3), 295–318.

Scherer, K. R. (1981). Speech and emotional states. In J. Darby (Ed.), *Speech Evaluation in Psychiatry*. New York: Grune and Stratton.

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

Scherer, K. R. (1992). Vocal affect expression as symptom, symbol, and appeal. In H. Papoušek, U. Jürgens, and M. Papoušek (Eds.), *Nonverbal vocal communication: comparative and developmental approaches*. Cambridge: Cambridge University Press.

Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of Voice 9*(3), 235–248.

Scherer, K. R. and J. Oshinsky (1997). Cue utilisation in emotion attribution from auditory stimuli. *Motivation and Emotion 1*, 331–46.

Schmuckler, M. (1989). Expectation in music: an investigation of melodic and harmonic processes. *Music Perception 7*(2), 119–193.

Schorske, C. E. (1979). *Fin-de-siècle Vienna: politics and culture*. London: Weidenfeld and Nicolson.

Schubert, E. (1998). Time series analysis of emotion in music. In *Proceedings of the 5th International Conference of Music Perception and Cognition*, Seoul, pp. 257–263.

Schubert, E. (2000). Regression modelling continuous data in music psychology. In S. W. Yi (Ed.), *Music, mind, and science*. Seoul: Seoul National University Press.

Schulkind, M. D., L. K. Hennis, and D. C. Rubin (1999). Music, emotion, and autobiographical memory: They're playing your song. *Memory and Cognition 27*(6), 948–955.

Scott, S., A. Young, A. Calder, D. Hellawell, J. Aggleton, and M. Johnson (1997). Impaired auditory recognition of fear and anger following bilateral amygdala lesions. *Nature 385*(6613), 254–257.

Sedlacek, K. and A. Sychra (1963). Die melodie als faktor des emotionellen ausdrucks [melody as a factor in emotional expression]. *Folia Phoniatrica 15*, 89–98.

Serafine, M. L. (1983). Cognition in music. *Cognition* 14(2), 119–183.

Serafine, M. L. (1988). *Music as Cognition: the development of thought in sound*. New York: Columbia University Press.

Shweder, R. A. (1994). You're not sick, you're just in love: Emotion as an interpretive system. In P. Ekman and R. J. Davidson (Eds.), *The Nature of Emotions*. New York: Oxford University Press.

Singer, J. L. (1970). Drives, affects, and daydreams: the adaptive role of spontaneous imagery of stimulus-independent mentation. In J. Antrobus (Ed.), *Cognition and Affect*. Boston,MA: Little, Brown and Company.

Singh, P. and I. Hirsh (1992). Influence of spectral locus and f0 changes on the pitch and timbre of complex tones. *Journal of the Acoustical Society of America 92*(5), 2650–2661.

Sloboda, J. A. (1991). Music structure and emotional response: some empirical findings. *Psychology of Music 19*, 110–120.

Sloboda, J. A. (1992). Empirical studies of emotional response to music. In M. Jones and S. Holleran (Eds.), *Cognitive Bases of Musical Communication*. Washington: American Psychological Association.

Sloboda, J. A. (1996). Emotional response to music: a review. Presented at the Nordic Acoustical Meeting, 12-14th June, Helsinki.

Sloboda, J. A. (1997). What do we know about what makes us respond strongly to music? context, content and performance. Paper presented at "Exploring Intense Musical Experience", a conference of the Society for Research in Psychology of Music and Music Education, 15th March 1997.

Sloboda, J. A. (2000). Everyday uses of music listening: a preliminary study. In S. Yi (Ed.), *Music, Mind and Science*. Seoul: Western Music Research Institute.

Sloboda, J. A. and P. Juslin (in press). *Music and Emotion: Theory and Research*. Oxford: Oxford University Press.

Smith, C. R. and A. Steinschneider (1975). Differential effects of prenatal rhythmic stimulation on neonatal arousal states. *Child Development 46*(2), 574–578.

Sperber, D. and D. Wilson (1995). *Relevance: Communication and Cognition* (Second ed.). Oxford: Blackwells.

Storr, A. (1992). *Music and the Mind*. London: Harper Collins.

Stratton, V. N. and A. H. Zalanowski (1989). The effects of music and paintings on mood. *Journal of Music Therapy 26*(1), 30–41.

Stratton, V. N. and A. H. Zalanowski (1991). The effects of music and cognition on mood. *Psychology of Music 19*(2), 121–127.

Stratton, V. N. and A. H. Zalanowski (1994). Affective impact of music vs. lyrics. *Empirical Studies of the Arts 12*(2), 173–184.

Strongman, K. and P. Russell (1986). Salience of emotion in recall. *Bulletin of the Psychonomic Society 24*(1), 25–27.

Sundberg, J. (1987). *The Science of the Singing Voice.* Illinois: Northern Illinois University Press.

Tembrock, G. (1975). Die erforschung des tierlichen stimmausdrucks [research into animal vocal expression]. In F. Trojan (Ed.), *Biophonetik.* Mannheim: Bibliographisches Institut.

Thayer, R. E. (1989). *The Biopsychology of Mood and Arousal.* New York: Oxford University Press.

Tillmann, B., E. Bigand, and M. Pineau (1998). Effects of global and local contexts on harmonic expectancy. *Music Perception 16*(1), 99–117.

Tischer, B. (1991). *Die vokale Kommunikation von Gefühlen [Vocal communication of feelings].* Ph. D. thesis, Universität München.

Trehub, S. E., A. M. Unyk, and L. J. Trainor (1993). Adults identify intant-directed music across cultures. *Infant Behavior and Development 16*(2), 193–211.

Treitler, L. (1980). History, criticism and beethoven's 9th symphony. *19th Century Music 3,* 193–210.

Treitler, L. (1984). What kind of story is history? *19th Century Music 7*(3), 363–373.

Treitler, L. (1993). Reflections on the communication of affect and idea through music. In S. Feder, R. L. Karmel, and G. H. Pollock (Eds.), *Psychoanalytic Explorations in Music.* Madison, Connecticut: International Universities Press.

Trevarthen, C. (1993). The self born in intersubjectivity: An infant communicating. In U. Neisser (Ed.), *The Perceived Self: Ecological and Interpersonal Sources of Self-knowledge.* New York: Cambridge University Press.

Tulving, E. (1993). What is episodic memory? *Current Perspectives in Psychological Science 2,* 67–70.

Urban, G. (1988). Ritual wailing in amerindian brazil. *American Anthropologist 90*(2), 385–400.

Vaneechoutte, M. and J. Skoyles (1998). The memetic origin of language: Modern humans as musical primates. *Journal of Memetics: Evolutionary Models of Information Transmission 2*, 84–117.

Velten, E. (1968). A laboratory task for the induction of mood states. *Behavior Research and Therapy 6*, 473–482.

Verba, C. (1993). *Music and the French Enlightenment: Reconstruction of a dialogue, 1750-1764*. Oxford: Clarendon Press.

Vitouch, O. (2000, August). When your ears set the stage: musical context effects in film perception. In C. Woods, G. Luck, R. Brochard, F. Seddon, and J. Sloboda (Eds.), *Sixth International Conference on Music Perception and Cognition*.

Wagner, H. (1989). The peripheral physiological differentiation of emotions. In H. Wagner and A. Manstead (Eds.), *Handbook of psychophysiology: emotion and social behavior*. Chichester: Wiley and Sons.

Wallin, N. L. (1991). *Biomusicology: neurophsysiological, neuropsychological, and evolutionary perspectives on the origins and purposes of music*. New York: Pendragon Press.

Walton, K. L. (1990). *Mimesis as Make-believe*. Cambridge: Harvard University Press.

Waterman, M. (1996). Emotional response to music: implicit effects in listeners and performers. *Psychology of Music 24*, 53–67.

Waterman, M. (1997). Intense by not emotional? the problems with describing experiences to music. Paper presented at "Exploring Intense Musical Experience", a conference of the Society for Research in Psychology of Music and Music Education, 15th March 1997.

Wedin, L. (1972). A multidimensional study of perceptual-emotional qualities in music. *Scandinavian Journal of Psychology 13*(1), 1–17.

Williams, C. and K. Stevens (1972). Emotions and speech: some acoustic correlates. *Journal of the Acoustical Society of America 52*, 1238–50.

Winer, B. (1971). *Statistical principles in experimental design* (Second ed.). New York: McGraw-Hill.

Wittgenstein, L. (1966). *Lectures and Conversations on Aesthetics, Psychology and Religious Belief.* Berkeley: University of California Press.

Yalch, R. F. (1993). Using store music for retail zoning: a field experiment. *Advances in Consumer Research 20*, 632–636.