Title

A Comparative Study of the Effects of Music on Emotional State in Normal Adults and those with High-functioning Autism

Thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy of the University of London

Rory Allen

Department of Psychology, Goldsmiths, University of London

February, 2010

Certification of originality

I certify that the work presented in this thesis is my own.

Signed ……………………………………….

Date …….. 24 February 2010………………..

Abstract

It has been assumed that the social deficits inherent in autism imply that individuals with the condition will be unable fully to appreciate the emotional content of music. My aim was to test this assumption, and to explore more widely the similarities and differences between the experience of music in the normal population and those with autism. My first study used musically-induced mood changes and a behavioural measure to show that mood music has measurable effects on cognitive processes in a control group. The second study focused on high-functioning adults on the autism spectrum, using semi-structured interviews to investigate the part that music plays in their everyday lives, concluding that autism is no bar to a full appreciation of the emotional uses of music, though also suggesting a degree of impoverishment in the language they use to describe the emotions. The final set of experiments compared control and autism groups directly, using physiological (GSR) measures of arousal together with self-report of the emotions evoked by a set of musical items. Standardized questionnaires were employed to measure alexithymia (difficulty in identifying and describing feelings) in individuals. Although the autism group experienced comparable levels of physiological arousal to music, they used fewer words than controls to describe their emotional responses, a difference which correlated strongly with their level of alexithymia. My results are consistent with the hypothesis that in autism, the basic physiological and emotional component of their reactivity to music is functioning normally, but their ability to translate these reactions into conventional emotional language is reduced, in line with their degree of alexithymia. These results suggest that the preserved ability of music to generate emotional arousal in autism may lead to clinical applications for the treatment of alexithymia in autism and other conditions.

Table of contents

List of figures

List of tables

Acknowledgements

Firstly, I wish to thank my supervisor, Dr Elisabeth Hill, for her unstinting support and help, given whenever I asked for it. Her outstandingly prompt responses and comments to numerous early drafts of the chapters of this thesis made the process of writing it much faster and easier than it would otherwise have been, and should serve as a model for how a supervisor should respond in these circumstances.

Thanks also go to my friends and colleagues in the Psychology Department at Goldsmiths, University of London, in particular to Rob Davis for creating the apparatus and software that made the GSR experiment possible, and to Maurice Douglas for helping me to set up in a new office when I moved to full-time study and for much other assistance.

Finally, I want to thank my wife, Dr Pam Heaton, for her constant encouragement and support in the task of carrying through and writing up this work.

Dedication

While conducting research into adults on the autism spectrum, I spent a great deal of time individually with a number of participants. My use of semi-structured questionnaires in one study involved long and sometimes discursive conversations, during which I began to understand the nature of the daily difficulties and trials with which practically all of them have had to deal. I came to admire their unfailing courage under the impact of struggle and unhappiness, and to appreciate their abilities, their persistence and their intellectual integrity. I hope that this thesis may do something to counter the myths, ignorance and misunderstanding under which they frequently suffer. Without the help of my participants, this work would have been impossible. This thesis is dedicated to them, in gratitude.

Language is like a cracked kettle on which we beat out tunes for bears to dance to, while all the time we long to move the stars to pity.

Gustave Flaubert

Chapter 1. Music, emotion and autism: current perspectives

There is great disorder under heaven, and the situation is excellent.

Mao Zedong

1.1 SUMMARY

This thesis deals with matters at the intersection of three areas of science that are themselves vigorous and rapidly developing foci of research, namely music, emotion and autism. Before engaging with these individual areas and their interrelationships, I will outline the structure of this introductory chapter and indicate the aspects with which I intend to deal before developing the main theme, in subsequent chapters, of my own experimental work on the topic.

The introduction will first examine the topic of autism and of research into the disorder, both in its historical aspects and in current areas, especially those involving empathy, that are most relevant to this thesis. It will then focus on emotions, and in particular on empathy, both in a wider setting and in its connections with autism. There will be a section on music, and finally an argument will be put forward to show that music is a promising candidate for probing a particular aspect of empathy that is thought to be impaired in autism.

1.2 AUTISM

1.2.1 Diagnostic classification

Autism is defined in the current edition of DSM in terms of a triad of impairments (American Psychiatric Association, 2000). These comprise deficiencies in social functioning, in the use of language for social communication, and in behaviour (known as “restricted and repetitive behaviour”, and including obsessive or pointless rituals, and excessive resistance to change). At present, DSM-IV-TR includes three subcategories of autism, autism *tout court,* Asperger Syndrome, and PDD-NOS (Pervasive Developmental Disorders, Not Otherwise Specified). Individuals with autism and

Asperger Syndrome exhibit all three of the triad of impairments, the distinction between them being defined, essentially, in terms of the date of the acquisition of functional language. If a person did not acquire language by the age of 3 years, they have autism, otherwise Asperger's. Occasionally the term "high-functioning autism" is used to refer to individuals in the normal range of verbal intelligence, a category which in practice is often very similar to that of individuals with Asperger syndrome under the definition given above, although conceptually, the definitions are distinct. The term PDD-NOS is used to classify individuals who have difficulties related to the autistic triad of impairments, but where the full criteria are not met, for example if only two out of the three deficits are clearly present.

The reality of the autism/Asperger distinction has been the subject of some debate, and it is rumoured that it is likely to be dropped in the next edition of DSM. For the purposes of this thesis, the controversy is largely irrelevant, though in what follows, reference to any meaningful diagnostic implications will be made where necessary. Unless otherwise stated, the term "autism spectrum disorder" (ASD) will be used to refer to the spectrum of conditions currently subsumed under both "autism" and "Asperger syndrome". Since all members of the clinical samples will have diagnoses of one of these two conditions, and do not include anyone with PDD-NOS, there will be no consideration here of what, if any, conclusions may be drawn with regard to the PDD-NOS category.

1.2.2 Early theories of ASD

Since Leo Kanner and Hans Asperger independently described the condition in children in the 1940s, knowledge of the nature and developmental trajectory of ASD has grown substantially. There was some confusion in the early years between ASD and infantile schizophrenia, but the latter term is now seldom used. Unfortunately, after ASD had been clearly identified as a separate illness, psychodynamic ideas formulated by Bruno Bettelheim (Bettelheim, 1967) attributed ASD to emotional unresponsiveness on the part of the mother. This so-called "refrigerator mother" theory was enthusiastically adopted by the psychiatric community. It led to the unfortunate situation where the parents of autistic children were blamed for their child's condition, thus adding guilt and mental anguish to the usual difficulties attendant on looking after a child with disability.

However, subsequent, more scientific investigations of the condition made it clear that ASD was a pervasive neurodevelopmental disorder, with a strong genetic component. As the evidence against psychodynamic explanations began to mount, it soon became clear that some more plausible explanation for the condition must be sought, outside the realms of Freudian mythopoeia.

1.2.3 Social cognition accounts of ASD

Such explanations were not long in coming, and there has indeed been almost an embarrassment of them. They have tended to fall into two categories: those which focus on cognitive abilities (or more generally, the lack of them) that are usually only evident at a comparatively late stage of development, and those which attempt to seek an explanation in terms of some low-level neurological process or fundamental neurological impairment which may perhaps be capable of detection quite early on. These theories are not necessarily mutually exclusive: on the one hand, cognitive theories may extend their territory to include neurological mechanisms, whereas on the other, theories based on abnormalities observed in the brain must perforce postulate ways in which these translate into cognitive deficits, given that autism is defined by its effect on observed behaviour. Nevertheless, a given theory will generally reveal an overall tendency to favour either cognitive or neurological aspects.

The cognitive theories deserve credit as being the first in the field, and as representing a brave attempt to bring some order to a confused picture. The first category to be considered consists of the “social cognition” accounts. The classical example of this is the “theory of mind” deficit hypothesis pioneered by Baron-Cohen and co-workers (Baron-Cohen et al., 1985; Baron-Cohen, 1989), which proposes that an inability to understand the various “orders of intentionality” in the minds of others is a fundamental characteristic of autism, and that this explains the spectrum of social and communication deficits. Deficits in theory of mind are typically tested using “false belief” tasks, which measure the extent to which an observer is capable of correctly understanding that other people can hold beliefs which the observer knows to be false. Baron-Cohen found that children with ASD typically performed much worse in false belief tasks than matched control groups, even when the controls included groups such as Down syndrome children with lower verbal and non-verbal mental ages.

The theory is more descriptive than explanatory, as it did not, at least in its original form, propose a mechanism by which the deficits in theory of mind might be caused. However, it has enjoyed an extraordinary degree of success since its original inception, and has played an important role in setting the boundaries for thinking about ASD for more than twenty years. One of its significant contributions in the early days was to provide a clear alternative approach to the understanding of ASD, which was independent of psychodynamic theory and which provided a theoretical framework in which further experimental work on ASD could be undertaken.

Baron-Cohen's theory was criticised by Klin, Volkmar and Sparrow (1992), who pointed out that identifiable differences in behaviour were observable in autism before the age by which even typically developing children develop a theory of mind of any degree of complexity, so that it was impossible for deficits in theory of mind *per se* to be an adequate mechanism for the development of social deficiencies. Partly in response to this and similar criticisms, Baron-Cohen later elaborated his theory (Baron-Cohen, 1995) into a form which, he claimed, avoided these problems. He proposed that theory of mind skills were built up out of simpler cognitive components, acquired throughout development, starting with an "intention detector" and an "eye direction detector" in the first few months. These modules alerted infants to goal-directed movements, and to whether another person's eye gaze was directed at them or at something else. This would be supplemented by a "shared attention mechanism" at around one year of age, leading to acquisition of a basic ability to share and understand the mental states of other people. These represented the foundation on which a fully developed theory of mind could later be constructed, whereas failure of these comparatively simple mechanisms to develop in early infancy might lead to the compromised theory of mind abilities seen in ASD.

Whilst this explanation might account for some of the features of ASD and be consistent with experimental data on infants, as yet no clear experimental evidence has been provided that these particular "detector" mechanisms exist as discrete modules, separately from other cognitive processes. Klin and co-workers proposed instead an "enactive mind" theory (Klin et al., 2003). This has features in common with Baron-Cohen's later approach (Baron-Cohen, 1995), in that it suggests that the social cognition deficits seen in autism are due to an abnormal developmental trajectory. However, it provides a possibly more fundamental explanation, in terms of impaired salience and attentional mechanisms in ASD. The theory claims that normal development in ASD is derailed early on by the reduced salience of social stimuli to the autistic infant. If a baby fails to orient to its mother's face, for example, the first and most basic channel of social interaction and learning will fail to open up at the proper time. Support for Klin's critique comes from research which suggests that some indications which predict the later development of autism may be manifested at a very young age, indeed as early as two to four months, at a time when bonding with the mother is crucial to the baby's development (Zwaigenbaum, 2005).

Criticism of Baron-Cohen's theory from another angle came from Hobson (1993), who was unconvinced by the idea that typically developing individuals understand the feelings of others through the information-processing model implied by Baron-Cohen's formulation of Theory of Mind. He considered that the empathic process instead involved a capacity to engage emotionally with other people in a much more direct way, and to grasp the nature of their emotional experience by, in some sense, sharing it with the other person.

This acknowledgement of the role of direct emotional experience in the process of understanding the mental states of others, was to some extent integrated with the earlier theory of mind model by Tager-Flusberg and Sullivan (2000). They suggested that theory of mind should be considered as having two components, a social-perceptual and a social-cognitive component, the former corresponding to the process considered by Hobson.

At this point, it should be observed that it is needlessly confusing to call the process described by these models as "theory of mind", the word "theory" carrying as it does connotations of exclusively cognitive processes. It seems more appropriate to term to process of understanding the mental, including the emotional, states of others as "empathy". Hobson's model was clearly and exclusively an emotional contagion theory of empathy, and the model of Tager-Flusberg and Sullivan, although formulated by them in terms of the theory of mind vocabulary then in use, was also in all but name a theory of empathy, though one allowing of a cognitive component.

Tager-Flusberg and Sullivan argued for the existence of these two separate components of empathy in their separate dissociation in different clinical populations. The social-perceptual aspects of empathy appeared to be spared in children with Williams syndrome, whereas social-cognitive aspects were considerably impaired. They found both functions to be impaired in autism. This model was adopted in its essentials by Baron-Cohen (2003), who, however, used slightly different terms to refer to these components: "affective empathy" for the social-perceptual process, and "cognitive empathy" for the social-cognitive aspect. Baron-Cohen, like Tager-Flusberg and Sullivan, found evidence for pervasive empathic impairments in autism.

Although Baron-Cohen still uses these empathy terms to describe theory of mind deficits in autism, his focus subsequently appears to have shifted away from theory of mind approaches as causal factors in autism. He currently favours the "extreme male brain" theory of autism (Baron-Cohen, 2002), whose content is essentially summarised in its title. This theory emphasises the "systematising" style often found in autism, making a link between this characteristic and maleness. The theory was presumably suggested by the fact that the great majority of individuals diagnosed with autism are male, and derives some plausibility from the fact that high levels of social and communication skills are traditionally associated with women rather than with men.

The extreme male brain hypothesis has spawned one more aetiologically specific offshoot, the "foetal androgen" theory of autism, according to which the neurological changes characteristic of autism (whatever these might be) are caused by elevated levels of foetal testosterone (Knickmeyer et al., 2008). Given the practical and ethical difficulties involved in testing this hypothesis, it is hardly surprising that no such test has yet been made, and pending any real evidence in its favour it must remain an interesting, if precarious conjecture.

One difficulty with taking social cognition accounts as the sole explanation of the disorder is that they fail to account for the varied but often very marked non-social abnormalities in autism, which include a tendency to restrictive or obsessive and repetitive behaviour patterns, hyper-sensitivity to novelty and change in the environment, and frequently, diminished IQ. They are, in addition, usually more descriptive than explanatory.

One cognitive theory which does appear to account rather well for the non-social deficits in autism is "weak central coherence" (WCC): see, for example, Frith and Happé (1994). This refers to a cognitive style which favours the piecemeal processing of incoming streams of information, at the expense of contextual meaning. In other words, the capacity to integrate perceptions or ideas globally is lacking, and the focus tends to be on details at the expense of the broader picture. The theory has certain attractions, but suffers from the same lack of explanatory power as the "theory of mind" paradigm. Furthermore, it appears increasingly to be in conflict with the empirical evidence, even as a descriptive account of characteristics found in autism.

Before proceeding to the evidence against, it should be noted that one piece of evidence in favour of Weak Central Coherence arises from the high prevalence in autism of savant syndrome, or the presence of profound social impairment together with some special or spared skill (Heaton & Wallace, 2004). Studies have found "special abilities" in no fewer than 9.8% of autism patients, compared with just 0.06% for other patients that were institutionalized for SLD (Hill, 1977; Rimland, 1978). At first sight, this high prevalence appears to be well explained by WCC theory. WCC's notion of a weakness in global functioning might explain how individuals tend to become obsessed with one particular topic to the exclusion of others, leading to familiarity and eventual mastery of it through the sheer amount of time spent studying the minute details of it.

However, since this theory was put forward, several items of evidence have suggested that autism may be associated not with impairments in global processing, but on the contrary, with enhancement of local processing. One such finding which is especially relevant in the present context is the evidence that global perception for music appears to be intact in autism (Heaton, 2005; Heaton et al., 2007). More fundamentally, a number of studies by a group led by Laurent Mottron in Montreal have led to a separate independent approach, best considered under a new sub-heading since the mechanisms which it postulates are of a fundamentally different kind from those put forward above.

1.2.4 Perceptual and neurological accounts of ASD

The first evidence supporting the new approach arose out of a case-study of an autistic savant draughtsman, "EC", who showed not simply a preserved, but an outstanding ability to reproduce the proportions of objects, clearly a global property. It was this (as they saw it) clear refutation of WCC that led Mottron and Burack (2001) to put forward their "Enhanced Perceptual Functioning" model. Mottron and Burack hypothesised that the apparent local bias in EC's drawing abilities was due, not to a global deficit, but rather to a lack of any intrinsic bias towards either local or global features. This led to an apparent focus on local features, simply because such features are generally far more numerous than global ones. It was as though the local features simply "outvoted" the global ones by force of numbers. In addition, they took issue with the term "cognitive style" coined by Happé (1999), proposing that the cognitive differences found in autism were due to a more profound difference in brain organisation than was suggested by the word "style".

What Mottron and Burack proposed, essentially, was that the differences in behaviour seen in autism could be explained in terms of corresponding differences in perceptual processes occurring at a comparatively low neurological level. Furthermore, these differences, operating from infancy, could affect social as well as other areas of functioning by interfering with the normal developmental process. They cited one specific example of how perceptual abnormalities in infants with autism might lead to regulatory behaviours which, almost incidentally, might severely compromise normal social development. It is well established that toddlers with autism often attend to an object by looking at it out of the corner of their eyes, rather than directly. Mottron and Burack explained this in terms of the hyper-sensitivity to perception of detail, and hypo-sensitivity to perception of movement, that they had found in adults with autism (and hypothesised also to be present in these toddlers). Lateral vision is known to be generally associated with a lower acuity for detail and better movement perception than foveal vision; therefore, the use of lateral vision by the toddlers might serve to down-regulate the otherwise painfully excessive perception of detail in an object of attention, and up-regulate the otherwise deficient perception of movement.

This explanation of the toddler phenomenon is at the very least highly ingenious, although it appears capable of at least one alternative explanation: that looking at an object obliquely, confines its image to just one of the (left and right) visual fields of view, and therefore allows it to be processed visually in a single hemisphere. According to one neurophysiological theory of autism, the condition is due to under-development of the corpus callosum compromising cross-talk between the hemispheres: this suggests that it might be easier to process an object in a single hemisphere rather than relying on coordination between the hemispheres. However, the details of the explanation are in a sense irrelevant. The point is that abnormal visual habits, however caused, would be liable to disrupt the normal fixation of an infant on, for example, the social cues provided by its mother's face and behaviour, leading to downstream problems with social understanding. To this extent, the mechanism put forward by Mottron and Burack might correctly explain the developmental failures which it is designed to address, even if Mottron and Burack's detailed analysis of the way in which this mechanism acts is not entirely accurate.

A further development of the EPF was suggested by a study comparing autism and control groups in a number of performance tasks which measured both global and local processing abilities (Caron et al., 2006). This showed the expected overall superiority of the autism group in local processing tasks such as the well established Block Design test. However, this superiority was exhibited only by a subgroup of the autism participants, namely those possessing a visuo-spatial peak; Caron et al. attributed this to the enhanced functioning and role of the primary visual cortex (V1) in this subgroup. On the other hand, there was no evidence for any inferiority of any of the autism group to the controls in the global processing tasks.

The EPF model has the merit of being able to explain a number of features of autism using a comparatively low-level causal mechanism, and suggests brain areas – those involved in low-level visual and perhaps auditory perception – in which the deficits characteristic of autism might be localised, thus indicating ways in which the theory might be tested empirically using imagery (though to the author's knowledge, such tests have not so far been conducted).

Another model which also has the merit of hypothesising specific brain areas as the site of abnormalities to account for the condition, and which indeed puts a neuro-physiological explanation at the forefront, is the executive dysfunction theory of autism (a recent description of which can be found in e.g. Hill, 2004). The executive functions, which include a range of such vital procedures as planning, working memory and inhibition, "have been linked to frontal structures of the brain, and to the prefrontal cortex in particular" and "share the need to disengage from the immediate environment in order to guide actions" (Hill, 2004, p. 191). Hill points out that the core deficits of rigidity and perseveration observed in patients with damage to the frontal lobes, could also account for many of the peculiarities observed in autism. If people with autism do indeed have abnormalities in this brain area, this could explain not only their social deficits (mental flexibility being essential in generating a fruitful social relationship) but also their "strong liking for repetitive behaviour and sometimes elaborate rituals" (Hill, 2004, p. 191).

Further evidence for this hypothesis is provided by the observation by Zwaigenbaum et al. (2005) that children who are subsequently found to exhibit autistic traits display, at a very early age, a reduced ability to disengage from a visual stimulus when a second stimulus is presented to one side of them. This "sticky fixation" is characteristic of typical infants at two months, but in infants who go on to develop autism it is still seen as late as four months, and appears to be an excellent predictor for the later development of the condition. Zwaigenbaum et al. suggest that "impairment in the disengage [sic] mechanism of visual attention may underlie the early social orienting deficits in autism, that is, the reduced tendency to orient and attend to ever-changing, novel and socially relevant stimuli. Disrupting the normal bias towards social orienting during the critical time when developing brain circuits are being guided by input from social experiences may set the stage for developmental pathways characteristic of autism" (p. 146). Zwaigenbaum et al. do not come down in favour of any particular neurological explanation for these phenomena. Despite citing other findings that: "Atypical patterns of cortical activation involving the prefrontal cortex have been observed in preschool children with autism" (p. 150), they do not refer to the executive function theory of autism, and do not even appear to be aware of it, though it surely provides the most parsimonious explanation of their own findings. Finally, some workers have attempted to make a link between the theory of mind and executive function theories of autism (for a recent example from among an extensive literature, see Pellicano, 2007).

Other researchers have focused on the neurological rather than the cognitive abnormalities found in autism. Just et al. (2007) observed both a smaller cross-section in the corpus callosum in autistic participants, and a correlated reduction in frontal-parietal functional connectivity, suggesting that not only may executive functions be compromised in autism, but that long-distance communication between brain areas may also be affected. Belmonte et al. (2004) claimed that deficits in connectivity and the ability to coordinate different areas of the brain might provide an explanatory framework for autism. Confusingly, others have found not a reduction but an increase in the size of the corpus callosum in autism (Lewis et al., 2005).

Other theories of autism have put forward different and additional candidates for the role of the neurological culprit in autism. Baron-Cohen et al. (2000) cite the amygdala as the guilty party, observing under-activation of the amygdala in autism during a task involving demands on the so-called “social brain”, in which the amygdala is a key component. This dovetails neatly, one might almost say too neatly, with his earlier “theory of mind” theory of autism, which centres on deficits in social behaviour. There is also a cerebellar theory of autism (Courchesne et al., 1994), which attempts to explain attention-shifting deficits in autism through problems with the cerebellum rather than the frontal lobes. Trumping them all, as one might say, some workers believe that the problem involves an overgrowth of the entire brain (Mosconi et al., 2006). It is indeed plausible that such an overgrowth, perhaps involving a partial failure of cortical neural pruning mechanisms, is likely to lead to a random disruption of normal communication patterns with deleterious implications for the functioning of any brain systems that require long-distance connectivity.

1.2.5 Environmental and genetic factors in ASD

This large number of potentially competing models suggests that there may be more than one cause for the condition. This conjecture is supported by the fact that autism appears to be a spectrum of disorders, rather than a clearly defined condition that individuals either do or do not have. High quality diagnostic tools such as the Autism Diagnostic Observation Schedule (ADOS: Lord et al., 1989; 2000), currently regarded as the gold standard for the quantitative measurement of autistic traits (at least in children), provides a continuous measure of social and communicative deficits, with cut-off levels for being autistic or on the spectrum (though the authors stress that ADOS alone is insufficient for diagnosis, as it does not test the presence of restricted or repetitive behaviours from the third member of the triad of impairments).

The search for the ultimate cause or causes for autism must, logically, focus on two possibilities: an early environmental influence or some genetically based abnormality. Although it now seems unlikely that the once popular culprit of MMR immunisation can be responsible, the case has been vigorously argued that autism may be caused by environmental factors such as heavy metal or other forms of pollution, though firm evidence implicating specific environmental causes is still lacking (Lathe, 2006). Focusing therefore on the alternative, the existence of a spectrum of severity of autism with no clear clustering of cases either side of some natural cut-off point suggests that it might be caused by multiple genetic influences acting synergistically, just as variations in adult height or IQ are thought to have a genetic component that is influenced by differences at multiple loci.

One of the principal investigators into the genetics of autism is Michael Rutter. In Rutter (2002), for example, he concluded, on the basis of large scale twin studies and other evidence, that genetic influences were extremely important in the liability to autism, and probably involved a number of interacting genes. Beaudet (2007) claimed that already, 10-20% of autism cases were traceable to an underlying genetic cause, and suggested that an additional mechanism, *de novo* mutations or epigenetic changes in the fertilized egg, together with an inherited genetic vulnerability, might account for many more.

1.2.6 Autism and alexithymia

Although logically this section belongs with my earlier survey of cognitive manifestations of ASD, the importance of alexithymia in the condition has only been recognised comparatively recently, and even now only by an – as yet – small minority of workers. This, together with the fact that alexithymia is a central theme of this thesis justifies its highlighting in a separate section.

The term was first coined by an immigrant to the US, Dr Peter Sifneos, in 1972, who derived it using two words from his native Greek: alexithymia literally means “being without words for emotions” (see Sifneos, 2000, for a historical survey of his contribution and an update of his views). Sifneos, who began his career squarely in the psychodynamic tradition, studying repression in patients with psychosomatic diseases, found that many of his clients showed an inability to find appropriate words to describe their emotions, as well as experiencing difficulties in distinguishing their feelings from bodily sensations of emotional arousal. Sifneos suggested that the limited emotional awareness and cognitive processing of affect seen in alexithymia, might lead to individuals focussing on and amplifying the somatic aspect of their emotions, thus playing a causal role in the development of their psychosomatic symptoms.

The concept of alexithymia has now been widely recognised in the clinical literature, where it has been observed in a range of disorders, including PTSD, anorexia and bulimia, major depressive disorder, panic disorder and among substance abusers. However, the experimental psychology community has been slower to adopt the term, and this is particularly true in the field of autism. The first paper to apply strict scientific methods to investigate alexithymia in autism found that it affected, to some degree, around 85% of the sample investigated (Hill et al., 2004). Even now, alexithymia is a topic almost completely ignored in publications on autism. In addition to the paper just cited, Berthoz and Hill (2005), Lombardo et al. (2007), Allen et al. (2009), Allen and Heaton (2010), Bird et al. (2010) and Allen (in press) are rare exceptions.

Two variants of alexithymia have been distinguished, described with admirable parsimony as types I and II (Bailey & Henry, 2007): it is the type II variety that is most commonly found in ASD (Hill et al., 2004). Individuals with type I fail to experience or describe emotions, whereas those with type II show a normal or high degree of conscious *awareness* of emotions but a reduced capacity to cognitively *appraise* them. In other words, the person with type II alexithymia may experience a relatively normal range of emotions, but will have a deficit in the ability to label or identify them or discriminate between their own internal emotional states. This type of alexithymia has been termed 'cognitive alexithymia' (as opposed to 'affective alexithymia', or type I).

Alexithymia is a significant cause of distress, for a number of reasons. Research comparing ASD and typical participants has shown that higher levels of alexithymia are associated with higher levels of depression in ASD. It is known that depression and alexithymia are correlated in the general population, though to date, group sizes have been too small to show such an effect within an ASD sample (Berthoz & Hill, 2005). However, it is plausible to suggest that there may be a causal link from alexithymia to depression and anxiety, given that sufferers are confused by the nature of their negative internal feelings, and are therefore unable to identify their causes.

Alexithymia also makes it difficult for sufferers to regulate their emotions: if an emotion cannot be labelled, it is harder to decide what to do about it. Simply giving an emotion a name and thereby assigning it a cause may lessen its negative effects. Spinoza believed that ‘an emotion which is a passion ceases to be a passion as soon as we form a clear and distinct idea of it’ (in Russell, 1961, p. 557). For Spinoza, unlabelled passions are both negative and destructive, and this labelling process diminishes the power of our passions to do us harm.

To the extent that a person’s internal emotional state influences their behaviour in social situations, alexithymia can also contribute to the core deficits in social functioning that are characteristic of ASD. This is because an inability to identify one’s own mood state means that one cannot make allowances for the effect of this internal state on one’s behaviour towards others. Moreover, even if in some cases people with autism retain some sensitivity to social situations and can respond to them with appropriate changes of mood (eg with anger in response to frustration or belittlement), an inability to cognitively label these mood changes may further increase levels of arousal, due to the anxiety arising out of this disability, resulting in inappropriate responses or damaging overreactions.

It may be observed that there is a certain parallelism between the division of alexithymia into two components, involving emotional unresponsiveness and cognitive difficulties, and the previously cited models of empathy which posit a two stage process. This parallelism will be explored and made explicit, and its implications for the empathic process in autism made clear, in section 1.3 below.

1.2.7 Conclusion: whither autism research?

While for many years, researchers have tended to search for a unified theory of autism, that is, one deficit that can explain the complete autism profile of strengths and ‘weaknesses’, it is now likely that multiple causes exist. Data in support of this has been identified in recent years from genetic studies. In particularly, Happé, Ronald and Plomin (2006) have suggested that the three components of the triad of impairments may vary independently of one another, and be under the control of separate genes. This, if true, requires us to consider a continuous three-dimensional, rather than one-dimensional space of symptoms. This finding suggests that the quest to understand autism requires something of a shift in emphasis. It will clearly be necessary to address certain key questions about interactions between independent traits and the influence of the environment on the development of diagnosed individuals. More radically, one might ask: what if any meaning can the term "autism" rationally be given in these circumstances? Should one fix upon some arbitrary definition of the condition, based upon behavioural criteria, and attempt to determine which possible permutations and combinations of genetic, environmental and developmental impacts might possibly result in that end state? Or should one focus more on specific neurological and functional impairments that might produce specific combinations of symptoms with a resemblance to aspects of autism as defined in DSM-IV, with the aim of classifying the condition in terms of discrete endophenotypes, each of which has a clear aetiology?

To sum up, despite six decades of research into autism, there is still no conclusive evidence as to its developmental and neurological basis. It is, however, widely acknowledged that autism and autism spectrum disorders are heterogeneous. This heterogeneity is manifested in extreme variability in the extent and severity of core diagnostic features as well as in intellectual and language skills. Thus whilst intellectually able, verbal individuals can obtain the same broad diagnosis as those without language and with low IQ scores, the presentation of social and communication skills are very different.

The lesson from this is that any conclusions about the condition known as "autism" must for the time being remain tentative and provisional. The nature of the sample of participants being studied in any empirical study may profoundly influence the results, and any generalisations that one may attempt to draw about people with the condition must be given the proviso that they may apply only to a particular sub-population.

1.3 EMOTION AND EMPATHY

1.3.1 Historical basis for theories of emotion

It will become clear in the following parts of this thesis that in obtaining the experimental data, no particular theory or model of the emotions was relied upon. This is because rather than basing the experimental stimuli on any pre-existing theory of emotion, a pilot group of participants were allowed to choose their own set of emotion words in describing their responses to music, which then formed the basis of subsequent wordlists from which participants could make choices to describe their own internal states.

However, in order to locate the experiments within a wider research context for readers' benefit, it is necessary at least to outline current theories of emotion so as to motivate both the empirical and speculative results and suggestions which follow.

Although the academic study of emotion goes back at least as far as the ancient Greeks, perhaps the main milestone in modern times was the publication of Darwin's monumental book "The expression of the emotions in man and animals" (Darwin, 1872). This approach has led, directly and with full acknowledgement, to Ekman's famous classification of what he believes to be the basic and cross-culturally universal emotions in humans: happiness, sadness, fear, anger, surprise and disgust (in, for example, Ekman & Friesen, 1986). It is essential to bear Darwin's work in mind when attempting to understand Ekman's contribution. The key word in Darwin's title is "expression". It seems that a study which avowedly focused on the expressive side of emotions, risks going further than the data warrant in founding a whole theory of the emotional substructure on it.

The origin of Ekman's theory in Darwin's work explains an otherwise rather puzzling fact about it. Emotions are usually thought of as being directed towards some external object. This is clearly true of four of Ekman's emotions: fear, anger, surprise and disgust. Each of these can be seen as triggered by new stimuli. But sadness and happiness have more of the nature of a mood than an emotion. It is possible to be sad, or happy, without necessarily knowing why – or indeed whether there is an external cause at all. The reason for Ekman's inclusion of these two emotions is that they result in two clearly defined facial expressions, as is the case with his other four emotions. In fact, his whole theory of emotion is founded on the fact that human facial expressions appear to be hard wired to reflect certain felt affective states, in a way which is demonstrably culturally universal.

However, the generalisation from the outward and visible sign of an emotion in the human face, to the inward and spiritual experience of it as a subjective feeling, is one that may not be justified, and it is debatable whether the radical nature of this generalisation has been sufficiently appreciated. In effect, Ekman has glossed over the difference which Darwin was subtle enough to appreciate, namely that between the expression of an emotion and the emotion itself. The first is akin to a single measurement, and the second, to the concept which it is measuring. The error in confounding the two is akin to the error of using a particular IQ test to measure intelligence, and then defining intelligence as "that which is measured by my IQ test". This point is raised only to make it clear that there may be hidden philosophical difficulties underlying apparently sound psychological theories of emotion, if the difference between a construct and its individual manifestations is not sufficiently borne in mind. These difficulties may return to haunt one with all kinds of confusions and apparent contradictions when faulty or incomplete theories of emotion are then applied to areas, such as musical emotions, outside the field in which the theories were originally developed.

Ekman's is far from the only theory of emotion that has been developed; dimensional models have also been put forward, notably a two-factor model involving axes of arousal and valency. However, most applications to emotion research in music use some variant of the Ekman model, possibly with a modified set of emotion states reflecting the fact that musically induced emotions typically involve more complex states such as nostalgia or wistfulness.

Aside from the descriptive, taxonomic side of emotion theory, there has been a vigorous research effort devoted to unravelling the neurological bases of emotions, investigating the more cognitive and causal aspects. Perhaps the first steps in this were taken by

William James, whose theory that emotions are essentially physical, or physiological in nature, has led to radically new ways of viewing them. In essence, his view can be boiled down to the statement that when one sees a bear, one does not run away because one is afraid, but one is afraid because one runs away (as cited in Le Doux, 1998). The modern way of putting this claim, as developed by workers such as Le Doux (1998), is that seeing a bear activates an automatic fast neural link from the thalamus to the amygdala, which produces a series of physiological and physical reactions that prepare us for a ‘flight’ response. The slower links from the thalamus to the cortex, and from the amygdala to the cortex, then ‘tell’ our conscious selves what is going on, and cause us to feel fear, and incidentally allow us to override the initial reaction if one becomes aware that the bear is not after all a threat: if, for example, one notices that the bear is a tame one on a chain and cannot launch an attack.

One important twist on this approach is the “cognitive labelling” theory of emotion (see eg Schachter & Singer, 1962). In a series of experiments, Schachter and Singer demonstrated that perceived emotions, while dependent to some extent on physiological changes, also involved a cognitive element depending on expectations. They induced changes in participants by injections of adrenalin, and were able, by suggesting different scenarios, to induce them to feel either negative (anxiety, fear) or positive (excitement) feelings. They suggested that one should distinguish the physiological component of emotions from the conscious perceptions and interpretations of them, which are described in what is now a generally accepted convention as “feelings”. The feelings one experiences are triggered by physiological changes, but are not determined exclusively by them. The process of ‘deciding’ what to feel is not a conscious one, but it is determined pre-consciously by the interaction between our existing beliefs and expectations, and the incoming physiological and sub-cortical ‘messages’. It is as though to some extent one constructs or even confabulates one’s emotions, sometimes inaccurately, on the basis of what one’s bodily reactions are telling one. (There is growing evidence, incidentally, from such areas as the study of change-blindness, that confabulation may play a greater part in people’s everyday lives than most of them would care to admit: one is reminded of Ibsen’s wry comment “if you take away the life lie from the average man, you take away his happiness as well”).

1.3.2 Emotion and empathy

Emotions do not of course exist in a social vacuum. They have evolved on account of selection pressure on organisms to develop mechanisms for coping to threats to the organism from the environment, including from other conspecifics. It is therefore of selective advantage to individuals to be able to obtain insights into the emotional states of others, if only as an indicator of their likely future behaviour. Using the term empathy in the broad sense to denote the process of gaining insights into the mental states of others, it is clear that the emotional components of these states will be of high significance in most situations. Whilst many theory of mind tasks do not involve an emotional component, the usual connotation of the term "empathy" has a narrower significance, in which an emotional component is mandatory if the term is to be used. In the words of a recent review article (Preston & De Waal, 2002, p. 4): "All forms of empathy involve some level of emotional contagion". This is the viewpoint to be taken in this thesis.

Some key insights into the general nature of empathy have come from studies of autism, as described in section 1.2.3 above. Hobson's insistence (Hobson, 1993) that narrowly cognitive theory of mind approaches were inadequate to explain how human beings really understand emotional states in others, represented a valuable corrective to early versions of the theory of mind based on deficits found in autism. Later versions of theory of mind incorporating Hobson's shared emotion model of empathy as one component, such as that of Tager-Flusberg and Sullivan, can only with difficulty be classed as pure theory of mind theories at all. The word "theory" conjures up an image of a rather cerebral process, hardly akin to the social-perceptual or affective empathy components of later theory of mind models. Rather than being models of theory of mind, these should more accurately be described as theories of empathy.

The nature of the affective component of empathy is, according to these theories, much more closely related to the general emotional system than it is to cognitive abilities such as language. This component includes the ability "to make on-line rapid judgements about people's mental state from their facial and body expressions" (Tager-Flusberg & Sullivan, 2000, p. 62). It is only the cognitive (or "social-cognitive") component which includes the more traditional aspects of theory of mind, involving "the conceptual understanding of the mind as a representational system" (Tager-Flusberg & Sullivan, 2000, p. 61). This involves the ability to solve false belief tasks, and interacts closely with other domains including language. The key point, however, ignoring semantic distinctions, is that empathy works on two levels, a rapid, possibly unconscious process which gives a "perceptual" view of another person's emotional state, followed by a slower, more cognitively based, process in which use is consciously made of models of behaviour based on stored experiences, language ("that person is jealous" etc) and the deduction of logical consequences.

Other studies of empathy have developed on a different empirical basis. One important area of ideas about empathy arose in ethology, or the study of animal behaviour, rather than from the study of neurodevelopmental disorders. Empathy has been seen in creatures as diverse as albino rats and chimpanzees. However, the broad conclusions of the two approaches appear to be not dissimilar. A review article by two researchers into animal behaviour (Preston & De Waal, 2002) concluded that empathy could be best accounted for by the Perception Action Model (PAM). This also involves a complex process involving a strong emotional component. The perception-action mechanism of which empathy is a part, links two fundamental types of behaviour, motor behaviour and emotional behaviour.

Given that this model is based on animal studies, and that (non-human) animals lack language, the model focuses more on the affective or social-perceptual aspects of empathy than the cognitive. In animals, an automatic mechanism for mirroring the emotional states of others may translate directly into action, if only of a kind that demonstates to the other individual that they are sharing that state (a situation also seen in human infants, who will respond to sounds of distress in another infant with distress of their own). Preston and De Waal argued that such a perception-action link is the cement that binds many animal societies together. They do, however, take a rather too narrow view of empathy when they claim (p. 17) that: "With a perception-action model of empathy, there is not empathy that is not projection, since you always use your own representation to understand the state of another". This may be true of animals, but the point appears less convincing when applied to humans: it is possible, for example, for a person to feel empathy for a person suffering from a mental state which they have never experienced, such as clinical depression. It is likely that in humans, the more complex two stage model is needed to describe empathy fully. Nevertheless, Preston and De Waal do provide useful evidence to support the existence of an initial, automatic response to emotional states in others, in which that state is in some way represented internally in the observer.

The most detailed examination of the evidence for the existence of this affective empathy process in humans has been provided in a review article by Goldman and Sripada (2005). They investigated a particular context for empathy, that in which individuals perceive emotion through visually examining the expression in another person's face. They favour the sort of automatic mechanism proposed by Hobson, but they also examine the possibility that empathy is a fundamentally cognitive process involving the sort of theory of mind processes found in Baron-Cohen's early models of mind-reading. They call the latter type of explanation "theory-theories", and the former, automatic variety of model "simulation theories".

Simulation theories involve a mechanism similar to that proposed by Preston and De Waal. The simulation process is based on the hypothesis that one perceives the emotion in a person's face by simulating (and generating) the same emotion in oneself at a subconscious level, a process possibly involving the mirror neuron system. In other words, empathy does involve actually experiencing the emotion expressed by the other person, but not – except in certain cases – feeling it in the same way or to the same extent, nor, of course, interpreting it in the same way. It is, as it were, a subdued version of that emotion, but projected onto the other person so that one *knows* what the emotion is (feeling a twinge of it, as it were, in oneself) but one consciously *sees* it as appearing in the other person. Empathy, in this reading, does not preclude one consciously feeling the opposite emotion: for example, a person may see their detested enemy suffer pain and empathise with the pain (therefore understanding the emotion being felt by their enemy) whilst rejoicing at it. Goldman and Sripada do not rule out the presence of a cognitive component of empathy, acting as a slower, second stage of the total empathy process and using the information provided by the simulation mechanism as cognitive input.

The weight of evidence, so far, seems to be heavily in favour of the presence of some kind of simulation or emotion contagion mechanism as the basis for the first, affective, stage of empathy. However, the case is not by any means finally made. The concept of simulation in empathy has been strongly criticised from the philosophical angle, for example in Gallagher (2007), and the decisive empirical data that would establish it beyond doubt is not yet available. Nevertheless, for the purposes of this thesis the affective-cognitive two stage model of empathy is, at least, plausible enough to serve as a hypothetical foundation on which subsequent investigations can proceed, reserving the right to test its assumptions as necessity requires and as opportunity provides.

1.3.3 Empathy and alexithymia: similarities

The observant reader may have noticed a parallelism between the binary structures of the affective-cognitive model of empathy, and the binary nature of the distinction between type I and type II alexithymia. The author believes that this is no accident. In fact, it is argued that it may be possible to learn about empathy from studying the similarities between these two processes.

This is somewhat complicated by the fact that alexithymia refers to failures in a mechanism rather than to the mechanism itself. For the sake of the discussion, let us use the term “emotional responsiveness” to describe the process whose *absence* causes type I alexithymia, and “emotional insight” to describe that whose *lack* is responsible for type II alexithymia. Affective empathy can then be seen to be composed of not one, but two processes: perception of an external emotional stimulus (another person undergoing an emotional state), followed by an autonomic, internal emotional reaction mirroring that state. In other words, a sensory/perceptual process is followed by a process of emotional responsiveness. Cognitive empathy is then not the final part of a two stage process beginning with affective empathy, but the outcome of a three stage process, involving perception, followed by emotional responsiveness, followed by emotional insight.

The benefit of seeing empathy in this light, is that it allows the two fields of empathy studies and research into alexithymia to be brought together in a fruitful cross-fertilisation. Alexithymia has hitherto focused on the use of a rather limited range of experimental tools, usually involving self-report questionnaires (which will be described in detail in following chapters). Empathy, on the other hand, has been widely studied in humans and animals, using self-report, behavioural and imaging methods of investigation. Support for the author's view that the topics are intimately connected, and evidence for the nature of the connection, is now beginning to appear. This evidence is as yet sparse, reflecting the general lack of interest in alexithymia in the research psychology community, but as an example of what can be done by an energetic and well-resourced team, a paper by Bird et al. (2010) provides clear empirical backing for the notion that the mechanisms compromised in alexithymia play a key mediating role in empathy. This hypothesis will inform much of what follows.

1.3.4 Empathy and alexithymia: differences and perceptual abnormalities

Hill et al. (2004) has been cited above as claiming that samples with ASD have type II but not type I alexithymia. If the linkage with empathy were a straightforward one, this would suggest that people with ASD would exhibit deficits in cognitive but not affective empathy. Yet there is evidence, cited above (in e.g. Tager-Flusberg & Sullivan, 2002), that this is not the case: deficits in both stages of empathy are found. The resolution of this paradox must lie in the presence of an initial, perceptual stage in empathy, which is also subject to individual variability and which accounts for the differences.

It is suggested that the reason for the affective empathic deficit in ASD is that the first, perceptual stage is compromised, while the second, emotional responsiveness stage is intact, as reflected in lack of type I alexithymia. Compromise of the first stage would result in an overall deficit in the total affective empathic process consisting of stage one followed by stage two. There is support for this hypothesis in the fact, known from eye-tracking and imaging studies Dalton et al., 2005; Schultz, 2005), that people with ASD process faces abnormally: they focus on the mouth area rather than the more emotionally salient eye region, for example, and the neurological processing of faces is also abnormal, involving reduced activation of the FFA (fusiform face area). Evidently, the sensory and early perceptual component of affective empathy is abnormal in ASD, at least in the socially most common context in which empathy is required, namely the identification of emotions in faces.

It should be noted that these sensory/perceptual abnormalities have been observed primarily in the visual domain. Unusual features have also been found in auditory studies into ASD, in particular in a series of papers comparing speech-processing in a sample of children with ASD and matched controls (Jarvinen-Pasley & Heaton, 2007, Jarvinen-Pasley, Pasley & Heaton, 2008, Jarvinen-Pasley et al., 2008). These showed that the perceptual processing of speech was intact or even enhanced in ASD, with the ASD group actually showing superior pitch-detection capabilities in speech compared with controls. The tendency to process linguistic as opposed to perceptual information (such as pitch) was, however, weaker in children with ASD. This has implications for difficulties in acquiring speech, in that the same word spoken by different people at different pitches can appear like different words to children with autism, making it hard to generalise across speakers. However, it suggests, equally, that any differences in auditory perception in ASD are likely to have a damaging impact on the perception of music, certainly to the same extent as the known deficits in the perception of faces.

One might, indeed, expect *a priori* that music perception would not be so vulnerable to disruption in ASD: music is primarily a phenomenon that unfolds over time, so the structure and meaning of music is presented in a sequence that is not open to the sort of failures of saliency that appear to lead to focusing on the mouth region of a face rather than the eyes. A visual scene may be scanned in an almost infinite number of different ways, but the auditory scene that is presented in music can be "scanned" in only one way, namely in the sequence in which the auditory events arrive at the ear. This suggests that if one is looking to isolate and measure the purely emotional processing components of empathy, namely those which are compromised in alexithymia, one should investigate the use of auditory stimuli. These are likely to avoid the sort of confounds due to processing abnormalities that may impair comparisons of ASD and control groups when tackling tasks measuring empathy, which typically rely on the visual modality.

1.3.5 Experimental testing of alexithymia in ASD

If that is the case, then it is evident that if one wishes to study alexithymia in autism empirically, using a methodology not totally reliant on self-report measures, it will be necessary to take certain precautions if the result is to have any validity. In particular, it will be essential to avoid using a paradigm that amounts simply to a test of empathy, as the latter will be hopelessly confounded by the associated perceptual disorders. It would be as unreasonable to try to measure inherent emotional responsiveness in an ASD sample by exposing them to a series of emotional faces, as it would be to measure art appreciation in someone with severe myopia by analysing their responses to the ceiling of the Sistine Chapel. It will be necessary to find a perceptual stimulus, in effect, which does not suffer from the known difficulties caused by perception of faces. One such possible stimulus, which at least has the plausible initial advantage of acting through a different sensory modality, is that of music. Before setting out the final structure and aims of this thesis, in which the use of music will play a crucial part, it is appropriate to deal briefly with this area in a broader context, to which the following section will be devoted.

1.4 MUSIC

1.4.1 The origins of music

If you had been living in North-Western New Mexico some 75 million years ago, you might have awoken each morning to the booming chorus of one Parasaurolophus calling to another Parasaurolophus across the misty equisetum marshes. These animals, which were equipped with a 1.5 metre long trombone-like structure in a crest on the skull, may have been the first on the planet deliberately to make something approximating to what one might describe as music, in the sense of creating and modulating the characteristic pitch of a sound for purposes of communication (Sandia National Laboratories, 1997).

The point of this observation is that when discussing the importance of music to humans, it is worth bearing in mind the antiquity of music's existence on earth, if only as a reminder that any investigation into its role and function that may be undertaken is limited in time as well as space, and that the conclusions of such an investigation are, to an even greater extent than in most of psychology, dependent on the culture, habits and environment of the populations being studied.

Figure 1: Mayan mural depicting musical instruments in a victory celebration, Bonampak, Mexico.

Music is one of the few aspects of human culture which is thought to be universal. The Mayans lacked knowledge of the wheel, the Incas lacked both the wheel and a written language, but they both in common with all civilisations – indeed, all human communities – appear to have either invented or incorporated some form of musical performance into their lives (see Figure 1). In the case of the only widely known exception, the Taliban, it appears likely that their aversion to music stems less from an innate insensitivity to music, than to something seen as an external cultural threat, and perhaps also something that can evoke emotion and a sense of community and action on behalf of individuals and groups. Their prohibitions included such a wide class of items that the presence of music hardly suggests a specific dislike or insensitivity to that art form. One such list comprised "pork, pig, pig oil, anything made from human hair, satellite dishes, cinematography, and equipment that produces the joy of music, pool tables, chess, masks, alcohol, tapes, computers, VCRs, television, anything that propagates sex and is full of music, wine, lobster, nail polish, firecrackers, statues, sewing catalogs, pictures, Christmas cards" (Waldman, 2001). Yet even the Taliban recorded "jihadi hymns" to encourage young men to join their cause (Mustafa, 2009).

The cultural universality of music has suggested to some observers that music is, like language, a hard-wired product of evolutionary selection. Apes have neither a complex spoken language nor anything approaching to a musical culture, nor are they capable – with the possible exception of gibbons – of either producing or appreciating music. Certain physical characteristics of humans which make language possible are also essential for that most basic of musical forms, singing. Without a raised larynx and sophisticated breath control, anything more structured than gibbon-like howling would be difficult. To that extent, music is certainly dependent on human evolutionary changes. On the other hand, the intense pleasure that listening to instrumental music gives us is hard to explain in evolutionary terms.

To date, the campaign over evolutionary origins has tended to polarise around two extreme views, among whose most vocal exponents are David Huron and Steven Pinker. Huron (2001) sees music as an evolutionary adaptation, whose development facilitated social bonding within early groups of humans, and whose benefits in terms of individual and kin survival fitness would have encouraged selection for both musical aptitude and musical appreciation. Others in the same camp (it is hard to avoid military metaphors in this contentious area), equally wedded to the idea of music's evolutionary origins, believe that it arose as a product of sexual selection, with outstanding musicians gaining in sexual attractiveness, a notion that might be caricatured as the "Jimi Hendrix theory". Steven Mithen (2009) argues that music preceded language, and evolved as a sort of proto-language before the appearance of *Homo sapiens*, and that music and language diverged as separate entities only after the appearance of modern humans. On the other side of this divide, Pinker (1999) has described music as "auditory cheesecake", and considers that not only is it not the product of evolution, but also that it is not even an essential part of human life, a view which puts him in a position rather closer to the Taliban than most of us would feel entirely comfortable with adopting.

At present the middle ground between these positions is, as is usual in No Man's Land, sparsely populated, but one divergent point of view has been put forward convincingly by someone whose experience and knowledge of the sciences of music and language is arguably at least the equal of any of those cited above. Aniruddh Patel (2008) has pointed out that the choice between the evolutionary and anti-evolutionary positions is in fact a false dichotomy. He argues that music shares some of the characteristics of writing and of the use of fire, in that both of these other aspects of human culture are neither evolutionary outcomes, nor, to use Pinker's dismissive term, useless “spandrels”. Although fire is perhaps one of the most ancient of human technologies and may indeed go back to the days of *Homo erectus,* writing is provably of recent origin, too recent to have been an outcome of classical Darwinian evolution. But human life in its current form is almost unimaginable without writing. Likewise, believes Patel, music has become an essential part of what it is to be human, without necessarily being the product of a biological process involving natural selection.

While in the present state of knowledge the question of music’s origins is probably not resolvable, Patel’s proposal deserves careful consideration, if only as a reminder that the truth is not always to be found among the disputants with the loudest voices. Moreover, the question, though still open, is important, because as I will discuss in more detail below, some of those who lean to the evolutionary view of music’s origins claim that its inception as an essentially social activity means that social awareness is necessary for its full appreciation. This has been used to derive very real and present – and potentially damaging – implications for the ability of people on the autism spectrum to appreciate music. The fact that this argument is, as I hope to demonstrate, misconceived, does not obviate the need to explain the background to it.

1.4.2 Music and emotion

There have been numerous studies which appear to demonstrate the ability of music to induce changes in affective state, though these findings have proved to be controversial. The area of music and emotion benefits from being of interest to three largely separate groups of investigators: music psychologists, musicologists, and philosophers of music. The latter two groups have tended to be divided between those who see music as expressing but not inducing emotion, and those who see it as primarily inducing emotion. Music psychologists have tended to see music as a valid inducer of emotion, though with the proviso that the set of ‘musical’ emotions, while overlapping with ‘normal’ emotions, are not co-extensive with them.

As evidence in favour of those who see music as *expressing* but not *inducing* emotion, it is clear from the personal experience of most people that it is possible to perceive that a composer intends to convey a certain emotion in a piece of music, without our necessarily feeling it ourselves. Music intended to be pleasant can in the wrong circumstances, for example if one is forced to listen to it while waiting for a phone to be answered, be merely irritating. Some philosophers of music (eg Zangwill, 2004) argue forcefully that this kind of secondary emotion is the only one that one can feel in response to music. It has been maintained that since emotion by definition must have an object towards which it is directed, and since music, being a succession of sound waves, is not an object, therefore one cannot feel emotion in response to music. This is perhaps the sort of logic that only a philosopher is able to appreciate. Nevertheless, philosophers and musicologists have performed a service by their insistence on the distinction between emotions *in* music, and emotions induced *by* music.

One easy response to the musicologists might be that both sides are right: sometimes one perceives emotion in the music, and sometimes one feels an emotion induced by the music, and sometimes both. However, if it is the task of psychologists to understand how the mind works, this statement seems unsatisfactory: it lacks explanatory force. Music psychologists, as opposed to music philosophers, appear almost universally to believe that music does induce emotions in the listener, but their explanations seem somewhat ad hoc. A mechanism which combines elements from both music psychology and music philosophy, that may suggest how it happens, is proposed below.

The most parsimonious account would be one that draws on mechanisms that are already known to exist in inducing emotions. Given that emotions are typically felt towards some external object or person, this suggests that one may look at whether music in some way represents an object or person. As it happens, an idea not too dissimilar to this is current in the modern philosophy of music. One of the people identified with this point of view is Jerrold Levinson, who has claimed (Levinson, 2006) that one perceives emotions in music as the expression of emotion by an imagined musical "persona". So far, this takes one little further in the direction of understanding induced emotions, since Levinson is considering emotion in the music, not in the listener.

But in view of the work of the empathy theorists cited above, this final step is relatively simple. If one perceives emotion in Levinson's "persona" through the same mechanism by which one sees emotion in real people, namely empathy, and if the simulation theory of empathy is correct, then one sees the emotion in music by simulating (and generating) the same emotion in oneself. This suggests that the debate between “emotion in” and “emotion evoked by” music may turn out to be a case of two bald men fighting over a comb. If the simulation theory of empathy is correct, *both* interpretations may be true. One perceives the emotions in the musical persona – or perhaps in multiple personae engaging in some kind of imagined drama or narrative – and experience the emotions oneself as part of the very *process* of perception. This idea, that empathic mechanisms have a key role in the induction of emotion by music, is supported by evidence set out in Juslin and Vastfjall (2008). These authors in fact describe six emotional induction mechanisms, the other five of which comprise brain stem reflexes, evaluative conditioning, visual imagery, episodic memory and musical expectancy. Of these two, brain stem reflexes and musical expectancy, in fact induce states of arousal rather than emotions, i.e. they induce components of emotions. Evaluative conditioning and episodic memory arouse emotions that are specific to a particular learning process related to the circumstances in which the music was heard in the lifetime of a particular person, and therefore do not form a part of any universal theory of emotion induction. Only emotional contagion can play a role in explaining a process in which fully fledged emotions are induced by music.

The explanation of the mechanism of emotional contagion by music as set out in pp. 565-566 of Juslin and Vastfjall (2008) refers to Preston and De Waal (2002), and has much in common with the approach of Hobson (1993). However, references in the two later papers to mirror neurons as providing an explanation of the phenomenon of emotional contagion, rely on the existence of a system whose presence in humans is as yet far from conclusively demonstrated. A recent review article (Welberg, 2010) on the subject of mirror neurons cites only one study (Mukamel et al., 2010) providing direct evidence for a mirror neuron system in humans. This study, a rare instance in which it was possible to record the activity of single neurons in human participants, showed that certain neurons fired both when participants observed a particular action and when they performed that action themselves. This is a slender foundation on which to base a theory of emotional contagion, and it appears preferable to examine the mechanisms of emotional contagion in response to music *de novo,* based on clear neurological and experimental evidence. This approach will be taken in chapter 2.

1.4.3 Music and ASD

Questions about how the various deficits observed in theory of mind, affective empathy and related processes vary within ASD and generalize to other domains, are still poorly understood. Among these is the issue of whether people with autism are capable of fully appreciating perhaps the most quintessentially emotional of all the art forms, music. Unfortunately this lack of empirical data has not served to limit speculation on the subject. For example, in a passage describing people with autism spectrum disorders (ASD), Levitin (2006, p. 253) states:

> Their ability to "read" the emotions of others is significantly impaired, and this typically extends to their utter inability to appreciate the aesthetic qualities of art and music. Although some people with ASD play music, and some of them have reached a high level of technical proficiency, they do not report being emotionally moved by music. Rather, the preliminary and largely anecdotal evidence is that they are attracted to the structure of music. Temple Grandin, a professor who is autistic, has written that she finds music "pretty" but that in general she just "doesn't get it" or understand why people react to it the way that they do.

An immediate question posed by this account is whether it really is possible to appreciate musical structure well enough to develop "a high level of technical proficiency" and at the same time to display an "utter inability to appreciate the aesthetic qualities of music". The early work of Meyer (1956) highlighted how closely music's aesthetic and structural properties are yoked, and more recently Huron (2006) has developed an elegant model of music listening that shows how emotional experiences can be generated from the listener's expectations and perceptions of musical structure. Given also that the cognitive mechanisms involved in perception of faces and music, though perhaps overlapping, are unlikely to be identical, the rationale for predicting that impairments will generalize across the two domains is very weak.

In contrast, deficits predicated on the basis of social and communication deficits are more plausible. Peretz (2001, p.127) has written:

> Music is above all seen as a function serving social ends. Because autistic individuals are hampered at this level, it seems reasonable to expect them to have failed to develop an emotionally responsive system for music. Music would have a different meaning for autistic individuals.

The value of this argument rests, however, upon the highly controversial question of whether music "evolved" in a biological, or a cultural sense. Is music really "above all … a function serving social ends"? Does music not also serve important intrapersonal functions? It is also important to consider the fact that many individuals with autism wish to establish social contacts, though perhaps lacking the aptitude to do this. It is certainly not the case that indifference or avoidance of others is a universal feature of autism. For those individuals whose handicaps result in social isolation, music may serve as a substitute for social contact, giving them the opportunity to consider themselves part of a wider community of music lovers. Music may also enable some individuals to learn about emotions. Later in this thesis empirical data will be presented to explore these suggestions further.

1.5 THESIS STRUCTURE AND AIMS

The aim of this thesis is to demonstrate the value of an experimental approach to the study of alexithymia in ASD. There are two key steps along the way to achieving this aim. The first one, which will be addressed in chapter 2, will be to show that music does indeed appear to be effective in stimulating the first stage of the empathy mechanism in normal controls, that involving induction of a congruent emotion, or the affective empathy response. This will show that music provides a valid experimental substitute for such stimuli as expressive faces, which are known to be problematic for other reasons in ASD. The second key step is to show that in the high-functioning ASD population, there are no obvious perceptual barriers to the emotional expressiveness of music, of an analogous nature to the perceptual barriers which are known to hinder emotional responses to emotional faces. The evidence for this second conclusion, which will be qualitative rather than quantitative, will be presented in chapter 3.

The essential logic of the thesis will therefore boil down to the following: firstly, emotional responses to music are of the same nature as emotional responses found in other, ecologically valid social-emotional situations, and musical stimuli can therefore be used experimentally as a proxy for human emotion in probing empathic responses; and secondly, that the lack of a deficit in the perceptual stage of the two-stage process of affective empathy triggered by the music in ASD, implies that alexithymia can be measured in ASD through measurement of the associated empathic responses, without confounds due to perceptual differences such may be introduced if attempts are made to measure empathy using visual cues.

The following two experimental chapters, 4 and 5, will test these assumptions by developing and using an experimental paradigm involving standard musical items. The emotional responses to these will be measured by both physiological and behavioural means, and the position of participants on the autism and alexithymia spectra will be separately assessed through the use of standard self-report questionnaires. It will be found that emotional responsiveness to music appears to be normal in ASD, as measured by skin conductance reactions. Secondly, the outcome of an emotion-naming behavioural task will be found to correlate strongly with an independent measure of emotional insight provided by one of the standard alexithymia questionnaires.

In chapter 6 it will be argued that this result, together with the lack of any group difference in basic emotional responsiveness, suggest that music can be used as a probe to test for the presence, and degree, of alexithymia in ASD, without the assumptions and vulnerabilities inherent in the self-report measures which to date, have been the only means of assessing alexithymia in this or any other clinical condition. As a bonus, the results, it will be argued, demonstrate that simplistic models of musical emotional deficits in ASD based loosely on broad-brush arguments involving the alleged social/evolutionary basis of music, are unsustainable.

Chapter 2. The validity of music as an empathic stimulus in a control group

My dear sir, in this world it is not so easy to settle these plain things. I have ever found your plain things the knottiest of all.

Hermann Melville (*Moby Dick*)

2.1 SUMMARY

The introductory chapter raised questions involving empathy, in particular surrounding the apparent deficiencies in empathy seen in autism, and the involvement of empathic mechanisms in the induction of emotions by music. This suggests the importance of empathy as a link between the concepts of emotion and music, and further indicates that the reality of musical emotion induction, whether empathic or otherwise, needs to be substantiated. The study described in this chapter has two aims: firstly to explore the existence of an effect of music in inducing mood, as measured by a behavioural test of judgment of mood in faces, and secondly to determine, having established its existence, which of the two main categories of hypothetical empathic mechanisms best explains the outcome. It will be shown that a clear effect of music on mood does exist, and that one explanation, the simulation model, appears to explain the direction and size of the effect better than the alternatives.

2.2 ABSTRACT

In order to test differing hypotheses about the mechanism of mood recognition in face perception, 30 participants were asked to judge the expression of a series of morphed sad/happy Ekman faces while undergoing musical mood induction procedures. An effect of mood on face judgement was observed, with sad mood induction increasing the number of faces identified as sad. The presence and direction of this effect are inconsistent with the 'theory of mind' and the 'generate and test' theories of face

processing, but accord with three alternative simulation theories. This is the first clear experimental demonstration of affect infusion in a non-evaluative judgement task.

2.3 INTRODUCTION

2.3.1 Musical mood induction and empathy

Before considering the question of detecting and measuring emotions induced by music, and the related issue of whether autism and control groups differ on these measurements, it is necessary to probe the nature of musical mood induction experimentally. One reason for this is that musicologists and philosophers of music have questioned whether music can induce genuine emotions, or whether descriptions of musically evoked emotions are really just metaphors for emotions seen in the music. They have suggested that in reality, the closest music comes to inducing emotion is in its ability to provide aesthetic and perhaps other forms of pleasure. They maintain that the perception of emotion in music, or the impression that emotions are being induced by music, is an illusion, born of the need for human beings to provide metaphorical descriptions of reality.

The interpretation of results in subsequent chapters does depend rather strongly on the issue of whether musically induced emotions really exist. When comparing autism and control groups on their emotional reactions to music, it is important to be clear whether the reported emotions are real or nothing more than metaphorical constructs. One of the purposes of the study to be described in this chapter is to undertake a relatively simple experiment in order to test whether music can induce emotions. Given that the question at issue is whether self-reports of musical mood induction can be trusted, it is clear that some other instrument than direct self-report must be used to investigate the question. A further motivation for conducting the experiment in the form described below, is that it throws some light on the mechanism for detecting emotions in others, or empathy, and as set out in chapter 1 it is the author's contention that the simulation theory of empathy, combined with Levinson's persona theory of musical expressiveness, may explain at least one of the mechanisms through which music induces emotion in the listener. The remainder of this introduction will be devoted to explaining the theoretical and experimental background to theories of empathy, as applied in particular to the visual modality, where individuals are engaging in reading the mental states of others in their faces.

2.3.2 Empathy as "mindreading"

To achieve a working level of social competence, humans have to be able to identify the mental states, including the mood states, of other people. One of the best recent reviews of the empirical and theoretical evidence for how this process may occur is Goldman and Sripada (2005). Their approach took the form of a traditional literature review, rather than a full-fledged meta-analysis; given the complexity of the problem and the diversity in the empirical data, the time is not yet ripe for the sort of multiple statistical analyses of comparable effect sizes that would be necessary in a meta-analysis.

Goldman and Sripada focused on the aspect of empathy that they termed "mindreading", which they defined as "the capacity to identify the mental states of others, for example, their beliefs, desires, intentions, goals, experiences, sensations and also emotion states".

They classified the published explanations of mindreading into two broad categories. The first category of approaches to mindreading maintains that an individual uses some kind of naïve or folk-psychology theory to infer mental states in another person, on the basis of that person's behaviour or other environmental cues. In some versions of this approach, the naïve psychological theory is part of a specialised pre-existing module in the brain, and in other versions, it may be acquired through some kind of domain-general learning process. But in any event, the process of mindreading is a more or less clear, cognitive process. It is essentially information-based.

Although Goldman and Sripada did not make this comparison, the image of the process which occurs to the author when he reads about these theories is of something not too dissimilar to what happens when he looks out of the window in the morning and decides on the basis of the appearance of the sky, wind strength and temperature, what the weather is likely to be like that day, before dressing accordingly. Goldman and Sripada referred to this group of theories in a convenient shorthand as being versions of "TT" or “theory-theory”, meaning presumably that they comprise theories according to which humans theorize about the mental states of others.

A second approach maintains that an individual carries out mindreading not in accordance with some kind of theorising, but through a process of “simulating, replicating or reproducing in his own mind the same state as the target’s, or attempting to do so”. Goldman and Sripada group together all explanations which are based on this sort of concept as examples of “simulation theory” or ST. The core idea of ST is that the observer attributes a mental state to the person they are looking at, after reproducing or “enacting” with themselves the state in question. In other words, the observer replicates the target’s mental state on the basis of what can be observed of the target (perhaps just their facial expression), and if the process yields a certain mental state as the result of this, the observer attributes that mental state to the target (it is not, by the way, necessary to suppose that this process of replication is carried out voluntarily or even consciously).

In order to make this problem more manageable, Goldman and Sripada (2005) focused on a particular mindreading task, that of attributing emotional states to others based on their facial expressions (face-based emotion recognition, FaBER). They examined the competing claims of the TT and ST explanations for this task, and concluded that on the basis of the existing, mainly clinical evidence, the ST models best fitted the data. The following is a summary of the evidence they provided for this; the actual full references have not been provided here, as they are only marginally relevant to this thesis and can all be found in Goldman and Sripada’s original paper.

In parenthesis, it should be noted at this point that the extensive literature on what might be called the “emotional contagion” theories of affective empathy cited in chapter 1 will not be drawn on here, for the reasons adduced in section 1.4.2. It should also be noted that the existence of an “emotional contagion” or “simulation” mechanism for affective empathy does not preclude the concurrent existence of a “theory theory” mechanism, such as that dubbed “cognitive empathy” by Baron-Cohen, acting at a later and more sophisticated stage of the empathic process. The “theory theory” concept being tested in this chapter, however, states specifically that the theory theory mechanism is all there is – that it is sufficient to account for the phenomena of empathy, without requiring an automatic, fast-track simulation mechanism. Gallagher (2007) puts the case even more strongly, arguing that a simulation process is not only unnecessary, but actually impossible. The simulation theory, in other words, is inclusive and is tolerant of the presence of other mechanisms, whereas the theory theory is exclusive and is intolerant of alternatives.

2.3.3 Evidence for ST vs. TT in Goldman and Sripada

Adolphs and others studied the case of a woman who had suffered bilateral destruction of the amygdala, an area of the brain which is central in the emotion of fear. When her judgements of the expressions in photographs of faces were compared with matched brain-damaged controls, it turned out that she was abnormally deficient, in her failure to recognize fear in faces. A similar patient studied by Sprengelmeyer was also deficient in facial fear recognition. A study of a sample of nine amygdala-damaged patients confirmed these findings, but found that whilst the effect on recognition of fear was most marked, recognition of other emotions, especially anger, was also somewhat abnormal.

This apparent link between a deficit in the ability to experience an emotion and the ability to recognise it in faces, extends to at least two other emotions, disgust and anger. A study by Calder and colleagues of an individual with damage to the insula and basal ganglia found a selective deficit in his ability to recognise the facial expression of disgust; the insula is suspected to be involved with the emotional experience of disgust. In the case of anger, the dopamine system is thought to be important in the processing of aggression, and experimental evidence shows that participants who have been given sulpiride, a dopamine antagonist, show selective impairments in the ability to recognise angry faces.

Further evidence comes from a study by Wicker et al. (2003), who found that in typical, non-brain damaged participants, neuroimagery showed that the same neural regions (the insula) were found to be active in both the experience of disgust, and the observation of disgust-expressive faces.

Goldman and Sripada concluded that the apparent link between the neural mechanisms involved in experiencing an emotion and those implicated in recognising the same emotion in others, strongly favoured ST over TT. The link between deficits in experiencing and recognising an emotion is easily understandable in the context of an ST approach to empathy: someone impaired in experiencing an emotion will naturally be unable to simulate a process which includes that emotion. A paired deficit in the cases studied by Adophs, for example, where a person was unable to experience or recognise fear, is straightforwardly predictable according to the simulationist approach. By contrast, there is no obvious reason to expect a paired deficit under TT. The existence of the paired deficits in, say, fear, would require that conceptual representations of fear, or the processing of theoretical information *about* fear, should occur in the same region of the brain that underlies the *experience* of fear, and similarly for the other pairings.

Goldman and Sripada went on to consider other arguments which might be used to salvage a TT explanation of the empirical data, but the interested reader is referred to their paper for further details, which are not relevant for present purposes.

In what follows incidentally, the term “mood” is used rather than “emotion”. If one adopts the rather strict definition of some writers (e.g. Konecni, 2003), one would only count as an emotion a mental state in which there was a clear and unambiguous cause or object, whereas the situation under consideration is one where observers arrive at an assessment of the mood state in a face without necessarily understanding the cause of that mood state. Although a rather laxer definition of the term “emotion” has been adopted in this thesis, the mental state changes involved in this particular experiment appear rather subtle, and the term “emotion” might therefore be misleading. In fact strictly speaking, Goldman and Sripada’s discussion applies to face-based “mood” recognition rather than “emotion” recognition, but even from a strict point of view it would perhaps be pedantic to object to their use of the FaBER abbreviation, where the “E” stands for “emotion”, and this term is retained in what follows.

2.3.4 Further discussion of the ST paradigm and its implications

Goldman and Sripada described four possible ST models. One, the generate-and-test model, involves the perceiver generating a hypothesised mood, and "enacting" it, or producing a facsimile of it, within herself. She allows this "pretend" emotion to have the effects which a genuine emotion of this kind would have, including activation of the neural circuits which would normally produce the relevant expression. If this expression (or neural correlate) matches the target face, the emotion is then attributed to the target. If it does not match, a different simulated emotion is produced, and so on in a recursive process until a match is achieved. At every stage, the hypothesised mood comes first, and the comparison with the target comes afterwards. There are, as Goldman and Sripada point out, technical objections to this model (it seems too complex to be consistent with rapid observed covert mimicry of observed facial expressions) aside from the fact that it is a mechanism requiring an additional stage to the other ST models, that of matching a self-produced expression with the target face. For these reasons, Goldman and Sripada considered this variant implausible.

The other three variants of the ST model are the reverse simulation model, the reverse simulation with 'as if' loop, and the unmediated resonance model: these are similar to one another, in that target simulation, at some level, precedes mood attribution, though the details of the models differ.

In reverse simulation, the observer of an emotion-expressive face begins by imitating that facial expression, probably in an attenuated manner. At this point, Goldman and Sripada refer to evidence that the mechanism by which an emotion in a person produces a corresponding facial expression, is partially two-way. Even if one does not feel an emotion, making oneself produce the appropriate expression can feed back into inducing, to some extent, the corresponding emotion, as reflected in the folk-psychology advice to smile through one's tears. So after imitating the observed expression, the observer has only to classify her own emotional state, induced by the process of imitation, as expressing the same state produced in herself. This might all happen at a level below that of conscious awareness.

The reverse simulation with 'as if' loop allows the link between the observation and the felt emotion to be achieved more simply than with the reverse simulation model outlined above, where it was necessary to use the observer's own facial expression as a mediating mechanism. The 'as if' loop would bypass the facial musculature; it would be a direct link between a visual representation of the target's expression and the observer's emotion state, allowing the observer to simulate the target's emotion directly.

The final model considered by Goldman and Sripada is the unmediated resonance model. To use their own description of the model: "the idea here is that observation of the target's face 'directly', without any mediation of the sorts posited by any of the first three models, triggers (sub-threshold) activation of the same neural substrate associated with the emotion in question….this would parallel findings of mirror-neuron matching systems found in monkeys and humans, in which internal action representations, normally associated with producing actions, are triggered during the observation of, or listening to, someone else's corresponding actions".

Goldman and Sripada, though favouring some form of ST explanation over the TT model, and despite their doubts concerning the first (generate-and-test) model of ST, do not come down as supporting any particular version of ST. They acknowledge that the empirical data is insufficient to do this. Their review of the possible explanations is, however, valuable in classifying them, both as between the TT and ST types, and within the ST stable, between those in which mood attribution precedes target simulation, and those in which the mechanism works in the reverse order. For present purposes, this dual distinction, between TT and ST, and between the two varieties of STs, will be sufficient.

Essentially, then, the ST mechanism for FaBER as summarised by Goldman and Sripada involves either generating moods internally and comparing them with the target, or modelling the target's appearance in some way, simulating this, generating the corresponding mood in oneself, then attributing it to the person being viewed. So the attributed mood is experienced by the observer, although possibly at a subthreshold level; the viewer is aware only of knowing or believing that the other person is experiencing a certain mood state, not of experiencing that same mood fleetingly themselves. Although the question of assessing moods in others appears at first sight to be a typically cognitive problem, this version of the ST puts the question into a wider context. If this hypothesis is correct, then a *cognitive* outcome (eg "that person is angry") is the consequence of a subconscious *mood* state, experienced internally but seen as a property of the other person (to borrow a word from psychodynamic theories, the mood would be "projected" onto the perceived face).

This suggests that consideration of this cognitive problem involves a wider issue: the interface between mood/emotion processes and cognition. In a recent review article, Lewis (2005) has categorised current theories of the interplay between mood/emotion and cognition into three: appraisal theories, functionalist theories and trait theories. Appraisal theories address the way in which perceptions and evaluations of a situation elicit a particular mood or emotional state. Functionalist approaches examine the problem from the opposite viewpoint, and investigate the effects of mood and emotion on attention, perception, memory retrieval and social judgement. Trait theories examine the influence of personality traits and clinical disorders on the cognition-emotion interface. Each of these approaches has its merits, as does the dynamic systems model advocated by Lewis, which incorporates features of all three.

Within this context, ST theories of empathy fit comfortably into a functionalist framework. They postulate an effect of mood on judgement, but go further than this: if this model is correct, then not only does mood influence judgement, but it is absolutely fundamental to it. Supposing that one gains knowledge of another's mood state by simulating that state oneself; then if the mechanism for generating that mood within oneself is absent for some reason, one's ability to form an accurate assessment of moods and emotions in others will be hamstrung, however intact one's other cognitive functions. This observation might account for the crippling deficits in the ability to empathise seen in otherwise high-functioning individuals on the autism spectrum.

Goldman and Sripada (2005), though concluding in favour of STs, fail to provide the kind of experimental evidence which alone can give confidence in the causal aspects of the ST model. The empirical data which they adduce includes only one strictly experimental design, in which Lawrence et al. (2002) found that the administration of a dopamine blocker impaired the face-based recognition of anger. However, they did not attempt to measure whether the same participants were impaired in their ability to *feel* the emotion of anger. Likewise, the experiment by Wicker et al. (2003) showing that the same brain regions (the left anterior insula and the right anterior cingulate cortex) were activated by the experience of a disgusting smell and by the perception of faces expressing disgust, while strongly suggesting that the two processes are functionally connected, does not amount to experimental proof of it. The author considered that an experimental and behavioural approach would widen the range of empirical data against which to judge the claimed superiority of ST over TT hypotheses.

STs postulate that FaBER is mediated by experience of a mood which is then attributed automatically to the target rather than felt separately; one might draw an analogy with colour perception: colour is experienced within the brain, but perceived as a property of the target object, so one could deceive the colour sense by wearing, say, rose-tinted spectacles. Emotional equivalents of wearing rose-tinted spectacles, which temporarily (and harmlessly) affect the underlying emotional cast of a person's mind, exist, and are known as mood-altering techniques. One would expect such techniques to affect FaBER judgements in the same way as tinted glasses affect colour perceptions. If, on the other hand, FaBER judgements were based on some variety of theory of mind, one would not expect a marginal change in underlying emotional tone to affect judgements of mood in others. This suggests a way of deciding between the two classes of theories: investigate whether or not the judgements of participants performing a FaBER task, when simultaneously subjected to mood-inducers, depend in any way on the nature of the mood being induced.

One of the earliest mood-altering techniques was the method of Velten sentences (Velten 1968), in which participants read a set of either elating, or depressing, sentences. Criticisms of the demand characteristics of this technique led to alternatives, including that of musical mood induction (Pignatiello, Camp, & Rasar, 1986). This has been implemented in different ways by different researchers, but a typical scenario is that a number of short musical items are piloted to find those which appear to reliably induce certain moods or emotions, and these are then presented in randomized order to experimental participants. Participants may be asked to report explicitly on their internal mood states in response to the music, or may be given other apparently unrelated tasks.

Whatever the experimental context, where it is necessary to include a mood induction element, the use of music thus seems the most acceptable technique for wide scale use, not least because there can be few ethical objections to the use of music for mood induction with vulnerable populations. A further advantage of musical mood induction techniques for our purposes was that because it is auditory, the technique could be applied at the same time as, and without interfering unduly with, a FaBER task, whereas the Velten and other visual techniques would require separate sessions of mood-induction, followed by the experimental task.

Experimental procedures involving musical mood induction coupled with FaBER tasks have in fact already been published, in David (1989) and in Bouhuys, Bloem, and Groothuis (1995), but in neither case are the results usable in the present discussion. David's main interest was in hemispheric asymmetries in the perception of happy/sad chimeric faces, and although he did find an effect of mood on identification of faces as happy or sad (with more faces identified as happy after happy mood induction), he acknowledged the possibility that "the transparent nature of the mood induction procedure [Velten sentences] exaggerated this effect in compliant subjects". Bouhuys, Bloem, and Groothuis (1995) were seeking to explore a hypothetical mechanism to explain the development of clinical depression, via a vicious circle effect of negative mood → negative perception → negative mood. They were not principally concerned with mechanisms of empathy in normal individuals; they used schematic and unrealistic face stimuli, and their normal controls may have been heavily influenced by demand characteristics: the elimination of over half their original sample who failed to report significant mood shift responses to their musical stimuli will likely have compounded this problem by selectively retaining precisely those participants who may have been most compliant to the evident wishes of the experimenters.

It should be mentioned here that two interesting experiments designed to assess the effect of mood on the perception of emotion in faces were conducted by Niedenthal et al. (2000) and (2001). Participants were presented with a 100-frame computer-driven "movie", in which the expression on a face morphed gradually from one expression to another (in the first paper, the expressions varied from happy to neutral or from sad to neutral, and in the second, from happy to sad and vice versa). Participants were instructed to stop the movie at the point at which they no longer perceived the original expression. Niedenthal et al. showed in the first experiment that participants in whom a happy mood had been induced, tended to detect a change in expression from happy to neutral at a later point of the movie than those in a sad mood, and those in a happy mood detected a change from sad to happy at an earlier point in a sad to neutral movie. Paradoxically, the direction of the effect in their second paper was reversed; so, for example, those in a happy mood detected the change of expression earlier in a happy to sad movie than those in a sad mood. Niedenthal et al. interpreted these results as supporting their theory that "individuals might detect changes in the facial expression of another person through the feedback, and perhaps change in subjective state, caused by facial mimicry" (Niedenthal et al., 2001).

My first observation on these experiments is that Niedenthal et al.'s hypothesis clearly belongs to the family of simulation theories as outlined by Goldman and Sripada. Mimicry is indeed a likely process for generating a subjective state similar to that of the target face, and the mirror neuron system immediately suggests itself as one way in which this might be effected. There is recent evidence from neuroimaging (Lenzi et al., 2009) that the mirror neuron system is actively engaged when a mother is interacting with her baby and imitating its expressions, and the neuroimaging study of Montgomery and Haxby (2008) supported this conclusion. However, these studies are not conclusive, partly because the imitation in Lenzi et al. was deliberate and the purported simulation mechanism when observing faces does not involve conscious imitation, and partly because the Mirror Neuron System is not yet so well established neuroanatomically that its activation in an imaging experiment can reliably be identified: Montgomery and Haxby rightly refer to it as "putative" at this stage of our knowledge.

My second point is that the procedure used by Niedenthal et al. is a somewhat counterintuitive way of measuring responses to faces. The reason for their very elaborate procedure – what one might call the "Dutch auction" approach – was that they were keen to avoid criticisms of previous similar experiments which had asked participants to use verbal labels in their responses to facial expressions. Niedenthal et al. (2000) reviewed criticisms of these studies which claimed that it was also possible to interpret their results in terms of response bias or verbal priming. By using a completely non-verbal procedure, Niedenthal et al. hoped to avoid this criticism.

However, in doing so they introduced a method of such complexity that it invites its own form of criticism on the grounds of over-elaboration. The point at which participants decided that "the face no longer expressed its original emotion" (Niedenthal et al., 2000) would clearly depend on a comparison between the currently observed face and the original face. Any observed effect would therefore be due to a differential judgement, not a simple judgement of emotion in one face. Such a judgement might be of interest in examining how people read actual emotions in ecologically valid situations, but it is clearly a higher order effect than a simple judgement of expression in one face, and without knowledge of first order effects it becomes extremely difficult to argue for a particular mechanism on the basis of second order effects. When one finds in addition that changing the film type from sad to neutral, to sad to happy, reverses the effect, one is into third order constructs, and the validity of the conclusions becomes even more suspect.

My third point, though it is not a criticism of Niedenthal et al.'s paradigm, is that in both the studies cited above they used mixed stimuli in their mood induction procedure, using a combination of films and music to enhance the effects that either would have had separately. In their 2001 paper they applied the music induction procedure alone to one third of the participants, but since the results for this subgroup were not reported separately, it is not possible to use them as a test of the music mood induction paradigm. This mixed procedure was consistent with their aim of measuring the effects of mood induction on facial judgements, because the mechanism of mood induction was not directly relevant to the hypothesis under test, but for present purposes, in which I needed to determine the effects of specifically musical mood induction, the results of these studies could not be applied directly.

I therefore considered it necessary to design an experiment specifically to test for the effect of musically induced mood on face judgement in normal participants. This design would use a realistic set of faces, simultaneous presentation of mood-altering music, and a task of sufficient complexity to reduce or eliminate the possibility of deliberate compliance with any demand characteristics of the procedure. It was decided that the simplest dimension of face variation should be coupled with the simplest variant of musical expression, that along the happy-sad axis.

The previous experimental results cited above have consistently suggested that the induction of positive mood in participants tends to bias participants towards perceiving positive affect in facial expressions. This result is relatively easy to understand in the light of the ST, or in Niedenthal et al.'s words, the "facial mimicry" hypothesis. Goldman and Sripada (2005) put the process involved in ST as follows: "[the attributor] tries to replicate a target's mental state by undergoing…the same or a similar mental process to the one the target undergoes. If, in her own case, the process yields mental state M as an output, she attributes M to the target".

In this study, the precise mechanism by which the attributor "undergoes a similar mental process to the target" is irrelevant. If it is assumed that the process outlined in the previous paragraph holds, it can be proved by a simple argument that any exogenous change in mood state in the attributor will bias the attributor towards judging a target face in the same direction as that of a mood state change.

To illustrate the principle, suppose the attributor A, experiencing a neutral mood state, is presented with a neutral target face T and judges that T is neutral: T = N, say. So A has attributed a mood state M = N to the target. But one knows (or at least, hypothesises) that "[the attributor] tries to replicate a target's mental state by undergoing…the same or a similar mental process to the one the target undergoes. If, in her own case, the process yields mental state M as an output, she attributes M to the target". Since A has attributed N to the target, therefore A must herself have a mental state of M = N. But A began with M = N. Therefore, exposure to T has in fact caused *no* internal mood change in A.

Now suppose that A's mood is externally altered towards the happy direction: M = H. A is again exposed to T and asked to judge the mood of the face. A will go through the same process as above, but (assuming that exposure to T has the same effect as before, ie causes no internal mood change in A), A will continue to have a mood of M = H. In other words, having undergone "the same or a similar mental process to the one the target undergoes", A perceives a mental state of M = H as the output. She therefore attributes M = H to the target.

It would be tedious to repeat this argument for every other possible combination of prior states, exogenously induced mood changes and posterior states attributed to a target face, so perhaps one additional example will be sufficient. Suppose A's initial mood is happy, and that she judges the target T to be neutral. One can therefore deduce that after observation of T, A's mood, initially happy, has become neutral (because one is assuming that A's observation of her internal mood M always results in an output in the form of a statement "T's mood state is M"). In that case, a combination of an initial mood of M = H and observation of T results in a final M = N. It is no more than a factual description of this observation to state that observation of T has caused A's mood to change in the direction of sadness. At this point an additional, though very plausible, assumption is required, namely that the mood change in A caused by her observing T does not depend on A's initial mood state. If it were found that the direction of change was different – if it were sometimes in the happy direction and sometimes in the sad direction – the present argument would fail. But if this assumption be granted, suppose that A's initial state were exogenously altered from happy to neutral, ie in the direction of making A sadder than before. So initially, M = N. By hypothesis, observation of T would change it in the same direction as previously: in the direction of sadness. Therefore, after the observation, M = S. And by the mechanism that has been postulated, A will therefore report T's expression as being sad. Therefore, a change in mood has again resulted in a congruent change in perceived emotion in the target face.

Having argued that the precise mechanism whereby the ST process is achieved is irrelevant to the experimental hypothesis, one may speculate about how it does in fact come about. It is at least plausible, given what is known about mirror neurons from the papers cited above and other numerous studies into the MNS, that the mechanism involves the attributor's MNS activating that part of the motor cortex involved with facial expression so as to match the facial expression of the target, or at least so as to generate a similar profile of activation of the set of the main facial muscles which are well-known to be key in facial expression such as the zygomatic and corrugator muscles. Clearly this does not necessarily involve an actual and visible change in the observer's own features, though this may happen. But the hypothesis is that it will at least lead to some activity in the pre-motor or motor areas, and possibly some minimal stimulation of the muscles themselves.

The fact that an individual's facial expressions are linked to her emotional experiences means that a link exists in at least one direction; the facial feedback hypothesis (as discussed and confirmed in, for example, the work of Strack et al., 1988) maintains that the link is in fact a two-way one, so that an exogenous change in expression can in fact generate a (possibly very small) congruent change in internal emotions. If the facial feedback hypothesis applies to the small change in expression induced by the MNS, this would provide an explanation for "the process [which] yields mental state M as an output". Even if the observation of the target produces no detectable activity in the facial muscles, but if the MNS does no more than activate the appropriate loci in the motor cortex to a level insufficient to produce such activity, the argument would still go through on one further assumption. If the facial feedback mechanism is mediated by activity in the motor cortex, as appears plausible, measurable activity in the facial muscles would not be required to trigger the facial feedback mechanism, but merely motor cortex activation; and it is precisely this sort of activation that is observed in experiments designed to detect the activity of the mirror neuron system. Therefore, observation of the target T, activation of the MNS to simulate the facial correlates of T (possibly confined to the motor cortex) and activation of the facial feedback system to generate the corresponding emotion exhibited by T, would complete the chain of events necessary to demonstrate the mechanism postulated by Goldman and Sripada under the general heading of the simulation theory.

At this point the theoretical discussion, which though perhaps of some interest is not directly relevant to the hypothesis under consideration, will be concluded. The hypothesis to be tested in this experiment is the simulation theory, in the form stated above, taken from Goldman and Sripada, namely that when judging the expression in a face, the attributor outputs an answer which is a function of her own mental state, and that since her mental state after seeing the face is partly determined by her mental state before seeing the face, therefore changing the prior mental state should result in a systematic bias in the posterior mental state, and therefore in the judgements of the mental state of the target faces. This hypothesis was operationalised by presenting participants with photographs of the same face exhibiting a range of expressions ranging from happy to sad, with participants being given the forced choice task, of deciding whether they were either happy or sad, with the prior mood state of the participant being altered by musically induced mood changes.

As will be evident from the description given below, the actual structure of the experiment departed slightly from this precise model. Rather than taking the mood induction procedure as a discrete initial process and then separately exposing participants to the face judgement task, the mood-influencing musical stimuli were presented concurrently with the task. This procedure has precedents in for example Niedenthal et al. (2000), where a mood initially induced using happy or sad films was maintained using classical music during the task itself (Niedenthal et al. refer in this paper to other paradigms in which a similar procedure was employed). The rationale for this was that even a successful initial mood induction would probably wane in effect over the period of each block of face judgements, lasting perhaps two minutes, given that the mood induction effect was not expected to be dramatic. Moreover, explicit exposure to happy or sad music in advance of the block would be more likely to signal the purpose of the stimulus to the participant, which would make it more likely that both response bias and verbal priming effects would be observed.

2.3.5 The "Mozart Effect", and possible effects on reaction times

The term "Mozart Effect" refers to a still controversial phenomenon first raised by Rauscher et al. (1995), who claimed that listening to a Mozart piano sonata produced significant short-term enhancement of spatial-temporal reasoning. They showed that this effect was not produced by either a taped short story, or by "repetitive" music, and that the effect was limited to reasoning ability and did not act through the improvement of short-term memory. A subsequent study (Rauscher et al., 1998) extended this finding to rats, who showed improved maze-learning ability after early exposure to a Mozart sonata. These results were criticised, however, by a number of researchers including Steele (2006), though some (eg Lemmer, 2008) found evidence supporting the claims of Rauscher and co-workers.

One implication of this research (specifically that reported in Rauscher et al., 1995, and Lemmer, 2008) was that music type not only influences performance on a cognitive task but also basic physiological variables correlated with arousal such as heart rate and blood pressure. This suggested that in the present study, performance and arousal measures might be affected by which music type (happy mood compared with sad mood) participants listened to. It was therefore decided to analyse not only any bias that might be introduced by music mood into face mood judgements, but also the effects, if any, of music mood on reaction times. It was predicted that happy music would be associated with higher levels of physiological arousal and cognitive alertness than sad music, and that responses would therefore be faster in the happy music condition than when listening to sad music.

2.4 THE PILOT STUDY

The first task was to select items of music which could reliably be used to induce happy and sad moods in listeners. A total of ten short instrumental items from the classical music repertoire were initially selected, on the basis that in the judgement of the experimenter, they divided into five each in the category of happy and sad. The items included a piano piece by Eric Satie, and extracts of symphonies by other composers including Mozart and Beethoven. These items were piloted with a small sample of judges, who were asked to rate each of them on a scale of 1 (sad) to 10 (happy), and the two items which registered the lowest and highest average scores were selected. These were Gymnopaedie: Lent et Grave by Satie (sad) and the beginning of the third movement from Beethoven's Seventh Symphony (happy). The length of the passages, at just over two minutes each, was chosen to be sufficient to allow participants to accomplish each block of the task while the music was playing.

As a manipulation check to confirm whether the items of music that were judged to be happy and sad would indeed have the appropriate effect in changing mood, a further pilot experiment was conducted in which participants were asked to record their mood changes by manipulating a pointer visible on the computer screen, with the appearance given in Figure 2.

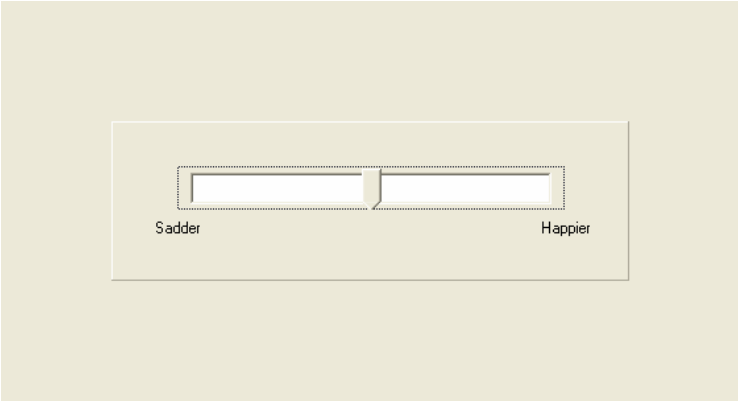


Figure 2: Appearance of computer-based mood indicator on the screen

Participants were asked, beginning with the pointer in the central position, to manipulate the pointer using the arrow keys on the computer to reflect the changes (if any) in their internal mood away from their mood at the beginning of the musical item. Participants were instructed that every time their mood changed while listening to the music, they should move the arrow by a corresponding degree in the appropriate direction.

Using a program written by Rob Davis of the Goldsmiths Psychology Department, the computer automatically presented the happy and sad items of music consecutively (in an order randomized between participants) and recorded the average position of the pointer during each item of music, using a value of zero on the extreme left (sad end) of the scale to 100 on the extreme right (happy end).

The items were trialled with a total of 15 pilot participants, recruited from staff and postgraduates in the Goldsmiths Psychology Department. The results were analysed using a paired samples t-test, with happy and sad moods as the two variables under comparison. The results gave mean ratings of 62.6 for the happy item and 48.8 for the sad item. The t-test yielded a value of $t = 4.2$, which was significant at $p = .001$. The manipulation check therefore confirmed that the internal mood changes associated with the happy and sad items were indeed present, and operated in the required directions.

2.5 THE MAIN STUDY: METHOD

2.5.1 Participants

Thirty university students, predominantly first year psychology undergraduates with ages ranging from 18 to 25, took part in the study (3 male, 27 female). They gave informed consent on the basis of an initial briefing prepared and approved by the Goldsmiths College Department of Psychology's Ethics Committee.

2.5.2 Design

The experiment was a repeated measures design; the two levels of the repeated measure were the valency of the audio stimuli, happy (H) or sad (S), which were played to participants while they carried out a forced choice task, that of deciding whether randomly presented morphed Ekman faces were on balance either happy or sad. The DV was the total number of faces in each block identified as sad, the aim being to see whether this would be affected by the valency of music.

2.5.3 Materials

The visual stimuli used in the experiment comprised nine female faces, which had been computer morphed in equal increments between standard Ekman faces for the emotions happy and sad. The faces were supplied by Professor Roberson (personal communication), based on a pair of happy and sad faces, of an individual identified only as "Belle", taken from Ekman and Friesen (1976), which were then cropped to remove external features and morphed using Morph 2.5 (Gryphon Software Corporation).

The picture in Appendix 2.2 shows faces 1, 5, and 9 in the series, corresponding to sad, neutral and happy.

Participants were asked to complete a short questionnaire about their musical education and preferences. This was intended to check whether there was any tendency of individuals with high or low levels of musical training and experience to show a greater or lesser effect of mood change on behaviour in the computer task. The questionnaire is given in Appendix 2.1. The questionnaire was adapted from one used to examine the musical experiences of children, in control and clinical groups (Downs, ASD), in

Heaton, Williams, Cummins and Happé (2007), principally by adding questions about current as opposed to childhood musical interests and activities.

2.5.4 Procedure

The visual stimuli (morphed faces) were presented on a computer, with the presentation controlled, and reaction times recorded, using e-prime. Participants were told that they would be shown a series of faces on the computer screen, preceded by a fixation cross, and that they should in each case react as quickly and accurately as possible to each face, to decide whether it was on balance happy or sad, and to press the H key for a "happy" judgement and the S key for a "sad" judgement. They were instructed that they must always choose either H or S, and in cases of doubt to go with their "gut feeling".

The experiment was initiated when the participant, after being given these instructions, pressed the space bar on the keyboard. A fixation cross was presented for one second, immediately followed by a face randomly selected from the morphed series, which remained on the screen for a maximum of ten seconds or until the participant reacted, when the cycle was repeated, with the fixation cross reappearing immediately after the previous keyboard response (in practice nobody took as long as ten seconds to decide on their reaction).

The stimuli were presented in blocks, each block consisting of an e-prime-generated random presentation without replacement of each of the nine morphed faces in four complete cycles. So for the first cycle, each of the nine morphed faces was presented exactly once in a random order, in the second cycle each face was presented exactly once but in a different randomized order, and this procedure was repeated twice more in each block. This meant that each face was presented exactly four times in each block, but the order in which the presentations occurred was randomized within each of the four cycles. For example, face 1 would always occur once during presentations 1 to 9, once during 10 to 18, once during 19 to 27 and once during 28 to 36, but the order of its appearance within these restrictions was arbitrary, according to a pseudo-randomizing procedure generated within e-prime.

The computer was programmed to pause at the end of each block to allow for manual changing of the music stimuli for the next block. The first block for all participants was presented without music, to allow participants to become used to the procedure. The subsequent two blocks were presented with musical stimuli in either SH or HS order, corresponding to happy and sad musical stimuli. The order of SH / HS blocks was counterbalanced across participants. Audio stimuli were played to participants at a standard, moderate volume on a Philips CD player with earphones. It was hypothesised that mood induction would only affect judgements on the more ambivalent faces, but the whole range was included to make the purpose of the experiment less obvious, and to provide a full range of RT data for subsequent analysis.

2.6 RESULTS

2.6.1. Face responses

The DV to be analysed for each level of the repeated measures IV was the number of faces out of the 36 stimuli presented in each block that elicited a press of the "S" key, ie that were identified as sad. The repeated measures IV had three levels, "no music", which was always the initial block for each participant, "happy", and "sad", in which the music of the corresponding moods was played.

The "no music" block was treated as a practice session, and the data for this block were not separately analysed. The means and s.d.s (standard deviations) for the remaining two blocks are shown in Table 1.

Table 1: Mean (SD) number of faces responded to as 'sad' (max.=36) for each type of mood inducing music, for the group and by subgroups (i.e. presentation order of music as happy, sad, HS; sad, happy, SH).

	HS group (N = 14)	SH group (N = 16)	Whole group
With happy music	19.5 (3.08)	19.2 (2.43)	19.3 (2.71)
With sad music	21 (3.28)	20.3 (3.00)	20.6 (3.10)

A repeated measures ANOVA with "mood" (made up of repeated measures Happy and Sad) as the within-subjects variable and group (order of presentation: HS or SH) as the between-subjects variable showed a significant effect of mood: $F(1, 28) = 13.43$, $p = .001$. The direction of the effect and the significant result meant that presentation of sad music, vs happy music, increased the number of faces identified as sad.

So as to be able to check for sequence and order effects (Mitchell & Jolley, 2007, pp. 411-418, and see also Discussion below), the main effect of group, and the mood by group interaction were also analyzed. There was no effect of group: $F(1, 28) = .241$, $p = .627$; and there was no mood by group interaction: $F(1, 28) = .274$, $p = .605$.

The APA recommends that p-values in results sections should be reported together with effect sizes, measured in some standard form, and with confidence intervals for the effect sizes (Wilkinson, 1999). In fact, measurements of effect sizes will turn out to be useful for other reasons, as will appear below. Cohen's *d* is one such standard effect size measure, applicable when an IV with two levels is under discussion, and is calculated by taking the difference between the means at the two levels of the IV and dividing by the overall standard deviation. Cohen's *d* provides a dimensionless measure which can be roughly judged as being small, medium or large for values of around 0.2, 0.5 and 0.8 respectively (Mitchell & Jolley, 2007, p. 312). In this instance the Cohen's *d* for the mood variable can be calculated from considering the difference between the number of faces identified as sad in the sad mood condition, minus the number identified as sad in the happy mood condition, for which $M = 1.30$, s.d. = 1.932; from which $d = 0.67$. SPSS yields an estimate for the 95% confidence interval of M as lying between 0.5784 and 2.0216, giving a 95% CI for Cohen's *d*, obtained by dividing these limits by the previous estimate of the s.d., of between 0.299 and 1.046.

2.6.2 RT data

It was predicted that RTs might be affected by music mood, with slower average RTs during the sad compared with happy music, so that if M_s is the mean RT for a participant in the sad music, and if M_h is the mean for happy music, then the new variable $M_s - M_h$, if calculated for each participant, would turn out to be significantly greater than zero, using a one-sample *t*-test. This difference was in fact non-significant

(t(29) = 1.2, p = .25), though the direction of the difference was indeed towards longer RTs with the sad music.

However, on examining the way in which reaction times varied with face stimulus presented for individual participants, it became evident that the variances in the RTs were considerably greater for the most ambivalent three faces in the middle of the range (numbers 4, 5 and 6) compared with the three saddest (1 to 3) and happiest (7 to 9) faces: the average within-subject s.d. for the central 3 faces was 378 ms, compared with 198 ms for the first three and 245 ms for the last three. The RT data for the central three faces for which the data were noisiest were therefore excluded, and a 2 x 6 repeated measures ANOVA with mood (sad, happy) as one within-subjects variable and faces (1, 2, 3, 7, 8, 9) as the other within-subjects variable was carried out. This gave a significant effect of mood on reaction time: $F(1, 29) = 4.461$, $p = .043$. The credibility of this result was supported by the fact that 22 out of 30 participants showed a longer average RT with sad than with happy music, and a non-parametric (binomial) test shows this difference to be significant, with $p = .016$ (2-tailed).

The same result can be examined more closely by calculating the difference between the average mean RTs for the six faces in each of the two mood conditions and conducting a within-subjects t-test on the result, which gives $t(29) = 2.112$, with $p = .043$ as before (as expected), and calculating the effect size as Cohen's d as above gives a value of d of 0.39, with a 95% confidence interval that the true value lies between 0.012 and 0.76.

In order to examine more closely the overall profiles of the RTs in response to different face stimuli and different musical moods, the average over all participants of RTs for each mood/face combination was plotted. A plot of the data for all nine faces in Figure 3 below, shows that the peak for slowest RTs appears to be shifted to the right for the sad mood music compared with the happy mood state.

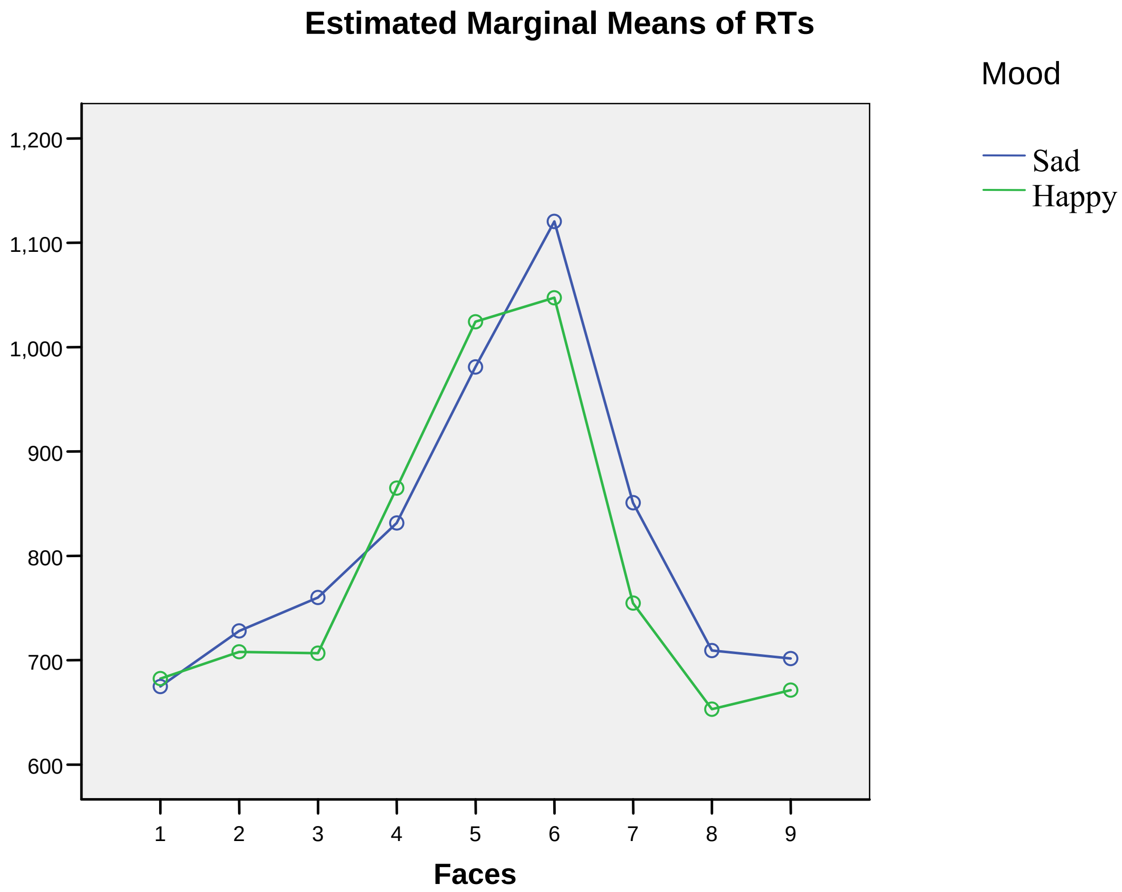


Figure 3: Average RTs (ms) as function of mood

There are two possible measures of the main effect of mood on individual participants. In calculating the effect size for the whole sample above, the difference between the number of faces identified as sad in the sad mood condition, and the number identified as sad in the happy mood condition, was taken as a measure of as the raw effect size, and this was divided it by the s.d. to obtain Cohen's *d*. For individuals, the same raw effect size was calculated by taking the number of faces identified as sad in the sad mood condition, minus the number identified as sad in the happy mood condition, for that participant. Another raw effect size measure for individuals is the difference for that participant between the average RTs in the sad mood condition and the happy mood condition. These two values, measured on the set of participants, can be correlated with one another.

There was no significant correlation between these two measures; $r = .029$, n.s. When the correlations were calculated between the level of musical experience encoded from the questionnaires, and the mood effect sizes above, the following results were obtained. The effect size based on sad face identifications correlated negatively and significantly with the level of experience, so that higher effect sizes were associated with lower levels of musical experience ($r = -.375$, $p = .049$). The correlation between experience and the RT effect size measure was not significant ($r = .304$; $p = .116$, n.s.). Note that two participants failed to complete musical experience questionnaires, so data involving these is limited to a sample size of N = 28.

2.7 DISCUSSION

The first analysis, of the RM ANOVA, confirmed that the effect of listening to sad music while conducting the face judgement task was to make it more likely that a face would be judged as sad, when compared with listening to happy music. This is the direction which was predicted on the basis of simulation theory. If the assessor judges the target mood on the basis of her own mood having seen the face, and if this mood is an arithmetic resultant of the mood of the target face and the musically induced mood, then sad music will clearly bias the assessor's mood, and therefore her judgement of the face, in the direction of sadness.

However, repeated measures designs such as this one are subject to validity threats, of which the most serious are sequence and order effects (Mitchell & Jolley, 2007, pp. 411-418). This was the reason for explicitly including in the analysis, the information as to which individuals belonged to which of the two counterbalanced groups, HS and SH. The presence of a significant main effect of group would have indicated the existence of a sequence effect, and the detection of a significant group by music-condition interaction would have suggested that an order effect had been encountered. The lack of a significant outcome for these results means that there is no reason to suspect the existence of either of these two threats to statistical validity.

The calculation of the effect size (Cohen's *d*) for the faces DV shows that the effect appears to be not only significant, but also not especially small, since the lower 95% confidence limit for it is 0.299.

The analysis of the RT data shows that the effect size of mood is lower than with the face-choice data, and that it is on the borderline of non-significance. Even this result was only achieved at the cost of removing the noisiest part of the data, and together this suggests that the previous analysis, of the face judgement data, is more reliable than the RT data.

Nevertheless, the RT data can help to provide at least some reassurance that they are consistent with the face judgement data and suggests an interpretation consistent with the simulation theory. There is no simple statistical method known to the author to analyse the apparent rightward shift of the RT peak in figure 1 under the sad mood condition. These results cannot therefore be presented as providing quantitative support for the main hypothesis. However, if viewed qualitatively, it can be shown that the general form of the profile has an easy and natural interpretation on the basis of the simulation hypothesis.

The face at which a peak is observed for a particular mood might be described as the "most ambivalent" face for that mood. It is clear *a priori* that the longest RTs shown by a participant, will be those for which the decision as to how to respond in the face judgement task is the most difficult. The decision will be the most difficult, when the face lies closest to the "neutral point" in terms of the listener's current standards of judgement, informed by her mood state. If she is listening to sad music, her judgement of the mood of a face will be biased towards the direction of judging the face as sad. Therefore, in order to be judged as neutral, a face will in fact have to be somewhat shifted in the direction of being *happier* than before. If she is in a happy mood, the face which is perceived as completely neutral and ambivalent will be shifted in the direction of *sadness.* This explains the observed differences between the peaks for the sad and happy music conditions.

One further point is that the profile provides some support in countering possible concerns over demand characteristics: it is very unlikely that deliberate hypothesisguessing could have been systematic enough to change the shape of the RT curve in precisely the expected way.

The analysis yielded two measures of effect size, that based on the actual face judgements, and that calculated from RTs in making the judgements. Logically, the former, being related directly to the task that participants were asked to complete, would appear the more fundamental figure. Indeed, the size of the first effect size appeared larger and more robust than the second. The fact that there was no significant correlation between the two effect sizes when calculated at an individual level may simply reflect the wide range in individual levels of responsiveness. It suggests also that one would not expect to find similarities in the correlations between these effect sizes and other variables at the level of individual participants. The negative correlation seen between level of musical experience and face judgement effect size, is therefore probably more meaningful than the small, non-significant one seen between RT effect size and level of experience. This correlation will be drawn upon in the discussion below.

The existence of an effect of mood on overall reaction times suggests that some kind of very basic "Mozart effect" may have been detected (Rauscher et al., 1995, and section 2.3.5 above). The lack of correlation between this and the effect of music mood on face perception tells us that there is no connection between two types of sensitivity to music, that which governs the effect of mood changes on FaBER, and that responsible for the effect of music on RTs. It seems likely that separate aspects of the music are responsible for these effects: the major vs. minor mode distinction may be responsible for the valency effect, and the differences in mean intensity or tempo for the RT effect. The experimenter's decision to use ecologically valid samples of music rather than synthesised items standardized for all parameters apart from the one under test, brings with it the drawback that possible confounding variables cannot easily be controlled for, and Bishop et al. (2009) have shown a clear effect of music intensity on subsequent choice reaction times in one recent experiment.

One possible objection to the design described in this chapter, which has been made in previous studies of a similar nature (see the remarks on Niedenthal's papers above), is that there might be a priming effect of the music on response, independent of the mood change. In other words, the brain might be becoming aware of the mood of the music, and could be being influenced by a simple cognitive link between the conscious awareness of the mood of the music, and the conscious act of button pressing to indicate the mood. This could provide a simpler and more parsimonious explanation for the results than a simulation-mediated mechanism.

However, if the effect were due to priming, one would not expect any correlation with musical experience, and certainly not a negative one. One would expect the conscious awareness of musical mood to be, if anything, stronger with experience. The negative correlation not only rules out priming as a confound, but also demands an explanation of its own. It is known that the brains of musically experienced individuals show more activation of the left-hemisphere language areas of the brain when listening to music (for example, Abecasis et al., 2009; Bever & Chiarello, 2009). This suggests that the listening experience is more under cognitive control than with inexperienced listeners, and one would expect less influence of subconscious or unconscious processes. On the other hand, it is postulated that the simulation mechanism is unconscious, and indeed can only work properly as an empathic mechanism if it is not allowed to intrude into conscious awareness.

This suggests that the empathic mood induction mechanism of music is likely to be suppressed or swamped by stronger conscious processes in the musically experienced. Musically more untrained or naïve individuals are precisely those in whom spontaneous emotional reactions to music are more likely to occur. Although knowledge of music may bring great benefits in terms of aesthetic awareness and sensibility, it may have a down side, in terms of reducing that "first fine careless rapture" which can be felt at first listening to a moving or an imposing item from the repertoire. There is some support for this view in Muller et al. (2009), who report that laypersons are "more inclined than experts to consider the ability of music to alter affective states as an important criterion for aesthetic judgements about music". If the affective content of music is seen as more important among laypersons, this suggests that they are more likely to be influenced by the affective content of music to which they are listening, while simultaneously carrying out a task such as having to judge the mood of faces, as in the present experiment.

The existence of a sizeable main effect of mood on face judgement, in the direction of sadder mood generating more sad face judgements, supports the ST model of FaBER, rather than the TT model which would predict a null result. The direction of the result, however, appears incompatible with the first of the ST variants, namely the generate-and-test model. Suppose for example that a participant was listening to sad music, and was presented with a neutral face stimulus. Under the generate-and-test hypothesis, the participant might first try to generate a happy mood internally, then check her resulting facial expression against that observed; if the expressions matched, she would find the face mood to be happy. But the saddening influence of the music would tend to counter her attempt to induce a happy internal mood, resulting in a neutral internal mood, which would produce a neutral expression in herself; since this matched the expression of the target face, she would conclude that the mood of the target was happy. Therefore under this hypothesis, sad music would bias face perception in the happy direction, and vice-versa.

In contrast the other three variants of the ST model, in which simulation by imitation of the target face preceded evaluation, would be subjected to an effect in the opposite direction, by the "rose-tinted spectacles" effect: sad music would bias internal mood and therefore judgements of face mood in the sad direction, and vice-versa. It seems likely, therefore that both the generate-and-test ST model and the TT model are incorrect. Rather, one of the three final variants of the ST model is correct, although it is not yet possible to distinguish between them using the data collected in the current study. This thesis is not primarily concerned with examining models of empathy in detail, so it is not proposed to devote space to examining possible paradigms that would distinguish between these mechanisms; it is likely that this would require a detailed technical examination of the temporal and spatial neurological correlates of the face-judgement process, in order to determine where, and in what order, brain areas are activated.

The fact that these results appear to exclude a TT explanation for the detection of mood in faces should not rule out a role for a theory of mind mechanism for the detection of emotion in the wider sense. If a target individual is experiencing an emotion rather than simply a mood, that emotion will have a cause or object. For an onlooker to arrive at knowledge of the target's emotional state, he or she must therefore arrive at conclusions about the likely object of the target's emotion, as well as about their underlying mood.

It seems plausible that this process of attaining this deeper knowledge of the target's emotion will happen on a longer time scale than the time taken in the current experiment to decide on and respond with a mood judgement (typically around 0.5 to 1 second), and will involve higher level cognitive processes which may be explicable in terms of a traditional theory of mind.

Whichever interpretation is put on these results in terms of deciding between alternative explanatory mechanisms, one of the clearest implications of this experiment is that listening to music of different moods has robust effects on behaviour and cognition in the listener. It also appears very likely that, at least for the simple case of happy and sad music, this effect is mediated by inducing corresponding happy and sad changes in mood in the listeners. Music does therefore appear genuinely to have great potency in altering mood states, though the negative correlation between the main effect size and degree of musical experience found in the present experiment, suggests that the degree of sensitivity to the indirect effects of music considerably affected by individual differences.

In subsequent chapters I will be focusing more intensely on this question of individual differences, and moving away from looking at the effects of specific musical moods on behaviour. In particular, I will, in chapter 3, explore in more depth the issue of whether high-functioning adults on the autism spectrum report being moved emotionally by music, and how they use it in their daily lives for mood manipulation and other aspects commonly reported in the music literature on the normal population. Later (chapters 4 and 5) I will investigate how individuals use language to describe their emotional and other reactions to music, and the extent to which individual differences in their responses correlate with other, quite separate measures of emotional fluency.

Chapter 3. Exploring the subjective experience of music, and emotional responsiveness to it, in ASD.[1]

I wish to propose for the reader's favourable consideration a doctrine which may, I fear, appear wildly paradoxical and subversive. The doctrine in question is this: that it is undesirable to believe a proposition when there is no ground whatever for supposing it true.

Bertrand Russell (*Sceptical Essays*)

3.1 SUMMARY

Having examined the nature of musical mood induction in a typical sample in chapter 2, this chapter moves on to look at the clinical population of interest in this thesis, that of high-functioning adults with autism or Asperger Syndrome. An examination of the literature on musical mood responsiveness in this group reveals a dearth of empirical data on the topic; studies on adults have tended to focus on gifted individuals, and experimental data on children with ASD cannot be extrapolated to the adult population. In view of the lack of prior examination of this topic, the alternative was adopted of a qualitative study, in order to map out the principal features of musical experience in an adult sample with ASD. This study revealed predominantly a range of similarities with previous studies of the TD population, but also one unforeseen difference: an emphasis on the use of internally focused descriptors of the emotional responses to music, suggesting possibly the presence of some restriction in the ability of this sample to verbalize their responses. The evidence for this conclusion is presented in this chapter, and explored further in subsequent chapters.

[1] The research in this chapter has been published as Allen, R., Hill, E., & Heaton, P. (2009). 'Hath charms to soothe ...' An exploratory study of how high-functioning adults with ASD experience music. *Autism, 13*(1), 21-41.

3.2 INTRODUCTION

In this chapter, the focus of research has moved from examining whether music can induce affective changes in a sample drawn from the normal population and demonstrating that these changes have real effects on behaviour, to investigating the nature of musical appreciation in a non-typical group, that of high functioning adults with autism. Hitherto, there has been little or no systematic research into the ways in which this group experiences music in the widest sense, that is, in the context of their everyday lives, a fact which has not prevented a number of speculative though surprisingly confident pronouncements on the subject. The aim of the work described here is therefore to provide a first, outline map of what is therefore a previously unexplored and uncharted country. Given that the priority is to cover the main and essential features over a wide area, rather than to focus on details contained within some arbitrarily narrow, clearly demarcated boundary, the approach adopted here – though commonplace enough in other areas of social science – will of necessity depart somewhat from the sort of clear cut statistical-experimental approach that is standard in psychology. The methodology that is most suitable at this exploratory stage of opening up a new topic is, as I will be arguing in more detail below, qualitative and synthetic, rather than quantitative and analytic.

Why should one bother to examine the area of music and autism in the first place? There are in fact several reasons why the nature of musical experience in autism should be of interest to psychologists. One is that music therapy has been used for many years in the clinical field for children with autism, with – it is claimed – considerable success. However, as is discussed more fully below, these claims have not been tested with the rigour that would serve to substantiate them in mainstream psychology; the question of whether the claimed effects are present is therefore at the same time unproven, and potentially of great practical significance.

3.2.1 Music and autism: conjectures and deductions

Another reason is that the issue of whether people with autism are "normal" in their appreciation of music is one that has been raised simultaneously in two areas of psychology somewhat removed from the study of autism itself: the evolutionary origins of music, and the study of Williams Syndrome. Williams Syndrome is, like autism, a pervasive neurodevelopmental disorder. As mentioned in chapter 1, it has been claimed (Huron, 2001) that the characteristics found in Williams syndrome, where "we see a group of individuals who are highly sociable and also highly musical", contrast markedly with individuals with autism, who display "extremely low sociability and also low musical understanding or affinity". Huron, who in this paper was primarily addressing the evolutionary origins of music, brought the Williams Syndrome/autism dichotomy into his argument to support his contention that the main evolutionary driver for music's origins was the role it played in social bonding. Given the DSM definition of autism in terms of a triad of impairments, including social impairments, it was useful for the logic of his case to argue that the social deficits in autism co-occurred with musical deficits. Williams Syndrome, on the other hand, though involving some of the same tendency to reduced IQ and other collateral neurological damage as is found in autism, has been described as preserving abilities in the social and musical domains (Levitin et al., 2005). This apparent double association between musical and social ability reinforced Huron's claims that musical abilities owe their origins, and indeed their current existence, to essentially social causes.

The claim that individuals with autism will be unable to appreciate music was repeated by another scientist working in the field of music research, when it was claimed in Peretz (2001) that: "Music is above all seen as a function serving social ends. Because autistic individuals are hampered at this level, it seems reasonable to expect them to have failed to develop an emotionally responsive system for music. Music would have a different meaning for autistic individuals … further investigation is worth undertaking". (p. 127). A similar claim was made by another researcher, as part of a discussion of autism: "Their ability to 'read' the emotions of others is significantly impaired, and this typically extends to their utter inability to appreciate the aesthetic qualities of art and music" (Levitin, 2006).

Though modern scientific practice generally requires the presentation of empirical findings before it allows us to proceed to generalisations, one may occasionally note a tendency for deduction to take the place of induction in the discourse of scientific claims. This has clearly been the case in this instance, where it has been argued on what appears superficially to be a basis of logic that (a) music is a social phenomenon, (b) people with autism are deficient in their understanding of social phenomena, therefore

(c) people with autism are deficient in their understanding of music. If this deduction were indeed merely the consequence of a form of argument valid in formal logic, and if the premises are true, then the conclusion must also be true. As far as the premises are concerned, a strong case has been made for (a), if it is interpreted to mean “music evolved as a consequence of its role in social bonding”. (b) is true by definition, given the definition of autism in DSM. Must one therefore accept (c), with the same flourish of conviction apparently shown by Huron, Peretz and Levitin? Or is it possible that their confident claims may be founded on something less than solid foundations? The approach of this chapter will be based on obtaining empirical data which may bear on the question.

3.2.2 Music and autism: empirical studies

Over the last two decades, an increasing number of experimental studies have explored an association between autism and absolute pitch, autism and savant skills in music, and autistic traits in musicians with absolute pitch (Bonnel et al., 2003; Heaton, 2003; Heaton, 2005; Heaton, Hermelin, & Pring, 1998; Heaton, Pring, & Hermelin, 1999; Heaton and Wallace, 2004; Heaton, Williams, Cummins, & Happé, 2007; Mottron, Peretz, Belleville, & Rouleau, 1999; Nettelbeck & Young, 1996; Treffert, 1988; Young & Nettelbeck, 1995). However, many of these studies report findings from pre-specified sub-groups of individuals, test specific components of music (e.g. absolute pitch) and do not therefore throw light on the nature of musical experience in the wider ASD population.

Experimental findings that generalize to broader populations of individuals with ASD have explored sensitivity to emotion in music. For example, Heaton, Hermelin and Pring (1999) showed that children with ASD understood the affective connotations of musical mode sufficiently well to be able to pair schematic representations of happy and sad faces with extracts of music in major and minor keys. In a more recent investigation (Heaton, Allen, Williams, Cummins, & Happé, 2008), typically developing 4-10 year old children, children and adolescents with Down syndrome as well as those with ASD matched musical extracts with pictures denoting a range of affective and non-affective scenarios; unimpaired performance was again noted in the ASD group. However, a limitation of both of these designs is that they specifically tested the ability to make conventional musical associations, and provided only limited insights into the nature of the participants' personal experience of music.

Other, non-experimental, investigations into the broader impact of music in people with autism have explored the value of music in a therapeutic context. Recent studies (Boso, Emanuele, Minazzi, Abbamonte, & Politi, 2007; Kern, Wolery, & Aldridge, 2007) have shown benefits, and a meta-analysis (Whipple, 2004) claimed that music therapy is demonstrably effective in improving the condition of children and adolescents with autism. For example, Wigram and Gold (2006, pp. 540-541) found some evidence of positive effect on interpersonal communication, reciprocity and relationship-building skills.

However, a number of methodological criticisms have been made of these studies (Accordino, Comer & Heller, 2007; Kaplan & Steele, 2005) and their conclusions should therefore be treated with some caution. The central points of these criticisms were that previous studies of music therapy had failed to ensure validity in three core areas. Firstly, there were no control groups, so that it was impossible to compare the therapy's effectiveness with improvements in a group of children not receiving therapy, which might also have shown progress due, for example, to natural maturation processes. Secondly, the evaluation of the children's progress was made either by interested parties (the therapists or their professional colleagues) or by parents for whom wishful thinking (and often the expenditure of large sums in fees) may have made the objective evaluation of progress difficult. And thirdly, the data were never subjected to a statistical analysis to ensure that the results reported might not simply be due to chance and/or selective reporting and inflation of type I error rates through multiple comparisons.

The present study will attempt to address some of these threats to validity by using a different approach. Qualitative analysis has different aims from quantitative, and the purpose of this study is not to show the existence of a 'treatment effect' as in the case of music therapy, but to attempt a relatively broad description of the features of the mental landscape of this sample of high-functioning adults with ASD, so far as that relates to their experience of music. The analysis is based on the words used by participants as the raw material, and so the sort of measurement of treatment benefits required to validate music therapy is not an issue.

Recent findings based on interviews and other self-report measures shed light on how typically developing (TD) people engage with music in their everyday lives (for example Batt-Rawden & DeNora, 2005; DeNora, 2002; Lamont, 2003; Lamont, 2009; North, Hargreaves & Hargreaves, 2004). These studies used a mixture of methods, some qualitative and some quantitative. Their results can therefore be compared with those of this qualitative analysis, and some comparisons made between ASD and normal populations, which may partly compensate for the lack of a control group in these interviews (the results of these studies will be summarised, and the comparisons made, and the appropriate part of the discussion section of this chapter).

These studies suggest that it would be fruitful to use alternative, non-experimental approaches to studying musical engagement in people with ASD, at least in cases where the well-documented semantic and pragmatic difficulties with language in autism (for example, Jolliffe & Baron-Cohen, 1999), do not make self-reporting of musical experience impracticable or unreliable. The study reported in the current chapter was designed to do this. I adopted a semi-structured interview format, aiming to enable participants to discuss the nature of their musical experiences with minimal guidance or interference from the interviewer. It was intended to limit the problems of unreliable reporting of musical experience, likely in those with relatively poor semantic and pragmatic language abilities, by confining the study to high functioning adult participants, and by using a semi-structured questionnaire with an open-ended interview approach.

Previous research has established that this methodology is appropriate for adults with high-functioning autism/Asperger syndrome (Hill, Sally & Frith, 2004). This approach has also produced some interesting results, for example in the hands of Batt-Rawden and DeNora (2005), when applied to samples from the general population. Batt-Rawden and DeNora made a qualitative study of 22 clinical individuals suffering from chronic illness, and claimed on the basis of interviews that a number of them had gained therapeutic benefits from music in terms of their general wellbeing.

Whilst an exploratory approach is unusual in psychology research, it has been recommended where an area has not previously been explored in depth. As Wilkinson (1999) states: “if exploration were not disguised in hypothetico-deductive language, then it might have the opportunity to influence subsequent research constructively”.

This study is not, therefore, cast in the usual form as designed to test a hypothesis formulated in advance and capable of making predictions which can be compared against actual data. The hypothesis, insofar as there was one, was simply that the semi-structured interview method would turn out to be fruitful, in terms of producing material that would throw light upon the actual experience of music in autism. If these participants had refused to discuss the topic, or had misunderstood the aim of my questions, or been unable to access the sort of personal memories and understandings needed to answer them, the study would have failed. Rather than inventing a hypothesis, this study was begun with a simple objective: to investigate the nature of the subjective experience of music in adults with ASD, exploring where the sample might be situated on the continuum of musical appreciation and understanding.

3.3 METHOD

3.3.1 Participants

Twelve high functioning adults (ten men and two women; ages 21 – 65, mean 41, s.d. 16) with diagnoses on the autism spectrum (nine with Asperger Syndrome (AS), three with autism) took part in the study.

Some of the individuals concerned had been given a relatively early diagnosis as children. Some, however, had only been diagnosed relatively late into adulthood. In several cases, individuals had, in adulthood, become aware of the core characteristics of autism from the media and had themselves insisted on being referred for a psychiatric diagnosis by their GP.

IQ data were not obtained for this sample. However, they were all high-functioning, in the sense of having no apparent language difficulties and being capable of clearly responding to interview questions. In all but two cases, participants attended appointments at Goldsmiths unaccompanied, indicating a baseline level of competence in day to day coping skills. In one case, I attended the participant's home for the interview, because he did not like to travel. This person had been a draughtsman designing electrical equipment, and had been made redundant six years previously, following which he had been diagnosed with Asperger Syndrome, which had enabled him to make a claim on his income insurance. He was by any standards, a high-functioning individual. One participant, a 21 year old man, attended the interview at Goldsmiths accompanied by his mother. He had been diagnosed with autism aged 14. Although less articulate, or at least, less talkative than the other participants, his intelligence did not seem to be seriously affected, as he illustrated by telling me an anecdote about the strategies that he had adopted for neutralising bullies when at school: strategies that involved a remarkably sophisticated theory of mind ability.

The design had been approved by the Goldsmiths College Department of Psychology's Ethics Committee, and all participants gave written informed consent. This was a convenience sample: participants had previously taken part in other autism-related research conducted by colleagues at Goldsmiths College on non-musical topics. Participants were not recruited for their musical abilities or experiences.

3.3.2 Materials and Methods

The first task was the development of the semi-structured questionnaire (text provided as Appendix 3.1). The research question involved exploring the nature of the musical experience in autism. As autism is a developmental disorder, our questionnaire focused on the early development, as well as the current nature, of participants' musical experience. This led logically to a division into two parts: "autobiographical memories" relating to music and "current reactions" to music. The specific questions in each section were suggested by the clinical experience of the author's supervisor, and by drawing on other music-related questionnaires. Certain questions were added in order to explore the possibility suggested by other authors that emotional deficits in autism would generalize to music. Two other questions, about the evocation of images and the importance of lyrics in songs, were suggested by spontaneous comments by two early participants mentioning these aspects. The addition of items to semi-structured questionnaires as research progresses, is acceptable in the Grounded Theory methodology within which the research was conducted (see below). Two examples of the texts of interviews are given in Appendices 3.2 and 3.3. These were chosen to demonstrate extreme cases, in one instance where the participant needed regular prompts chosen from the list of questions, and in the other, where the participant was so willing to talk unprompted on relevant matters, that he was allowed to comment at length on music-related topics of interest to him.

Responses to the questionnaire were analysed using a qualitative data tool, NVivo7 (QSR, 2006). The method of analysis was modelled on Grounded Theory, as developed from the work of Glaser and Strauss (1967) and based in particular on the methods for applying grounded theory to psychological research described in Pidgeon and Henwood (1997). It was felt that a qualitative approach was well-suited to the exploratory nature of the primary question: "how do high functioning adults with ASD experience music?", as per Losch (2006) and Pidgeon and Henwood (1997).

The method of analysis was divided into the phases of initial analysis and coding, core analysis involving the subdivision of the initial categories, and the formation of "outcomes": relationships, definitions and models (Pidgeon and Henwood, 1997). Bazeley (2007) suggests asking questions of the text to generate nodes: who, what, when, why etc. In the initial analysis of the raw transcripts, two clear categories emerged, dealing broadly with the questions "why" and "what". The term 'motivations' was chosen to represent the reasons or aims explaining "why" individuals decided to listen to music, and 'characteristics' to represent the type of music which was chosen to achieve those effects ("what"). All transcribed interviews were initially coded into two corresponding tree nodes; one containing all 'motivations' passages from the interviews, and one including all the 'characteristics' passages, each node subdivided by participant (for details of how the coding process works in NVivo, see Gibbs, 2002, or Bazeley, 2007).

I then separately examined these motivations and characteristics nodes, searching for themes within these subcategories that were common to more than one individual, creating sub-nodes for these emerging themes. This process of forming sub-nodes can be carried out electronically in NVivo7; it involves becoming thoroughly familiar with the raw transcript material, noting what appear to be common themes among two or more individuals, encoding the material according to these themes, and by an iterative process, reexamining the coded material and raw transcripts to see whether the themes appear to be consistent with one another and to divide up the material in a way that appears meaningful. The process ends when the categorisations appear to be stable and include all material of significant interest in the raw transcripts. Considerations of inter-rater reliability are not strictly applicable to this process, since it depends on a qualitative rather than a quantitative approach. The analysis involves an intensive time-consuming process on the part of the researcher, and it is not usually practicable to arrange for this work to be checked, at least in the context of the sort of approach that is necessary in conducting research at this level.

Details of the definitions finally adopted for inclusion in nodes, and an extensive set of quotes showing typical participant comments within each node, are given below at Appendix 3.4.

In order to establish/evaluate whether early musical training and experience were relevant to individual reports of the role of music in their lives, and the types of music that might be important to them, I also administered a questionnaire assessing levels of musical experience, based on a child version previously validated by Heaton, Williams, Cummins, and Happé (2007), and mentioned in chapter 2 above: see Appendix 2.1. The Musical Experience Questionnaire sets out questions relating to the formal and classroom experience of learning and performing music in childhood, and is designed to measure the extent of such experience and the degree of spontaneous liking of, and reactions to music in the early years. Each question is presented in a closed-question format.

3.3.3 Procedure

Participants took part in the study in the Department of Psychology at Goldsmiths College. In order to accustom participants to the interview situation I administered the early musical experience questionnaire first, reading out the questions and noting participants' replies. This non-threatening procedure served to break the ice at the outset of the interview and accustomed participants to talking about their musical interests.
I then proceeded to the semi-structured interview, recording this part (with participants' consent) on a mini-disc machine for subsequent full transcription. There was no fixed length to the interviews; they continued for as long as participants continued to provide

relevant responses. The actual lengths varied from around 30 minutes to roughly two hours.

3.4 RESULTS

3.4.1 Development of musical interests

Responses to The Musical Experience questionnaires as well as the interviews suggested that with a couple of exceptions, the participants fell into two categories, with five individuals in each. The first, which is referred to as the 'classical' group, acquired a liking for music at a relatively young age, and now enjoyed listening to, and in some cases also playing, primarily classical music. The second, 'pop' group, comprises those whose interest in music blossomed when they discovered pop music in their teenage years, having previously shown little or no interest in music. Two participants did not fit into either group, and were 'outliers' in a qualitative sense. Of these, one was interested in music only as an accompaniment to films; the other was interested mainly in the narrative structure and dramatic qualities of Wagner's operas. These two had both been diagnosed with autism rather than Asperger Syndrome (the remaining participant with autism was a skilled performer on the piano, and belonged to the classical group).

3.4.2 Why did participants engage with music: Motivations

The full tree diagram for the motivations nodes referred to in this thesis is included as Figure 4 below.

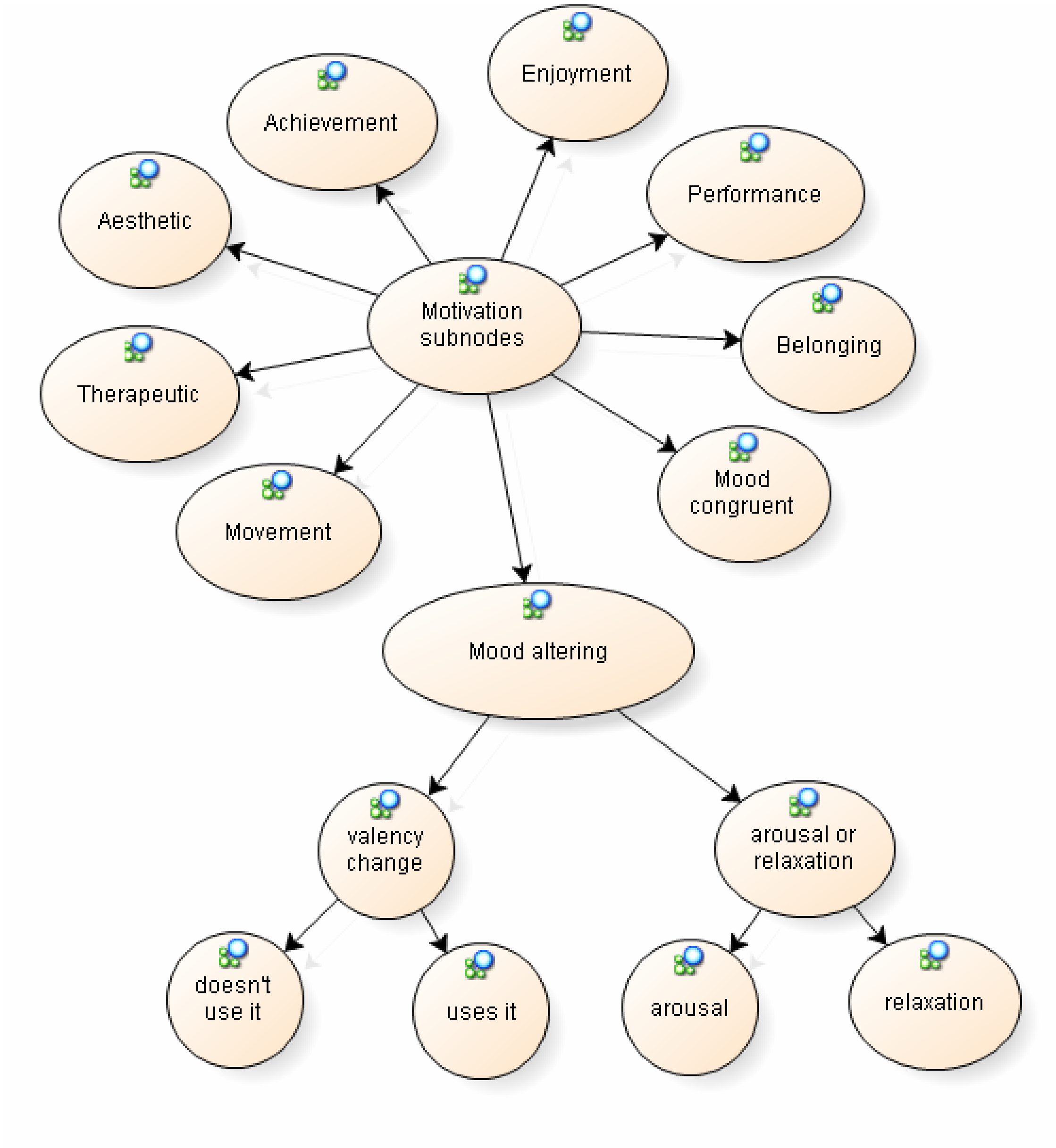


Figure 4: 'Motivations' subnodes

Participants used music extensively to achieve a number of personal aims. The largest number (nine individuals, or 75%) could be described as 'mood altering'. This node was identified early on, in our analysis of the motivations super-node. The words or phrases used to describe the moods aimed for, included: 'a buzz', 'exciting', 'ethos and excitement', 'exhilaration' 'feeling better about things' 'chillout', 'calmness or serenity', 'makes me feel at peace', 'relaxing', 'calming'.

The majority of this set of words can be grouped into two clusters. ‘A buzz’, ‘exciting’, ‘ethos and excitement’ etc are synonyms or near-synonyms for a state indicating pleasurable arousal. ‘Chillout’, ‘relaxing’ etc are synonyms or near-synonyms for states of pleasurable lack of arousal. Under the tree node for mood-altering, I created a subnode comprising statements about changes in arousal, with these two nodes under it to accommodate the two clusters.

To accommodate the three examples which did not fit clearly into these categories (‘cheer up’, ‘feeling better about things’, ‘sometimes made happy’), I created a separate subnode of ‘mood altering’, a subnode I denoted ‘valency change’. This subnode was coded into two further subnodes to signify both the (occasional) use of music to achieve an explicit state of happiness, and statements by a different set of individuals that they did not use music in that way. This is illustrated in the model included at Figure 4.

Though it is not reflected explicitly in the coding, the motivation for achieving a more relaxed frame of mind was in several cases stated as the need to counter anxiety or tension: “if you feel tense and you want something to calm you down”; “so that I don’t get over-anxious and over-heated up”; “when I’ve felt disappointed [i.e. frustrated]”; “music is calming if I’m anxious, it distracts me from unpleasant thoughts in my head”.

However, one person explicitly denied using music to change his mood in any way. Two out of the three instances of the use of music for valency change (i.e. to achieve change of mood along the sad/happy axis) were qualified by the participants, for example: “When I started getting more severe medical problems a few years ago, then I was playing happier music to cheer myself up. Not exactly cheer me up, but keep me on a level.” And five participants explicitly denied using music to induce valency changes.

After “mood altering”, the next most numerous category was “Aesthetic appreciation”, with seven (58%) individuals mentioning this. The “aesthetic” category refers, by the definition adopted for the purposes of this study, to feelings of pleasure that are derived from the conscious or intellectual perception of certain qualities of the music, rather than being a spontaneous emotional reaction to it. “Aesthetic pleasure” thus refers to a reaction in the listener; it is important to retain the distinction between this and, for example, “structure”, the latter being a characteristic of the music that might provoke a reaction of aesthetic pleasure, rather than that reaction itself. Included in the Aesthetic category, were participants’ descriptions of the pleasure they derived from the lyrics, the ‘ethos’, the pattern, structural complexity, or simply the ‘sheer beauty’ of the music.

The next commonest category was “Therapeutic”. Six participants (50%) reported experiences in which listening to music had had a healing effect. “Therapeutic” and “mood altering” categories both referred to changes of internal states. Confusion between the two categories was avoided by using “mood altering” to describe a cluster of statements comprising a limited range of words synonymous with either excitement/exhilaration, or calmness/relaxation, and containing no explicit statement suggesting a healing or therapeutic effect. A different cluster of statements described the use of music to heal emotional pain or depression, and these were denoted as “therapeutic”, for example: “When I have been feeling depressed, I have listened to certain music, and I would claim the music healed me”, and “Nothing worked for me - therapy didn’t work - but listening to this chap [singer Carl McCoy] actually worked”.

The next most popular headings, with five participants (%) each, were Belonging and Performance. These tended to co-occur, with four participants reporting in both categories. In the Belonging category, characteristic quotes were as follows:
“When someone’s released a classic album and you listen to it, I think ‘I’m part of something great now’”.
“I hear a lot of music in church, I go on Sunday and Monday and Thursday as a rule … you can relate to it, sometimes if you're not at that church, if the same song comes along sung by a different congregation, you can relate, it makes you feel a part of something…”.
“There’s something about pop music, it’s the image, and to do with people of my generation, rather than adults inflicting it on you.”

Two quotes illustrating different aspects of the Performance node were as follows:
“[as a child] I used to go into a nearby wood and sing at the top of my voice, which was exhilarating”.

“I’m a lot of the time a restless character, and I got a bit carried away with Queen, because I wanted to perform, to perform on stage and have people adoring me, I just wanted attention, I wanted people to say ‘oh look at me’, I’m a showman”.

Mood congruent, Movement, and Enjoyment, were chosen by four participants (33%) each.

“Achievement”, with three individuals (25%), included the two highly competent classical musicians in the sample and referred to their delight in mastering a difficult piece or for one individual, a difficult musical mode (e.g., Dorian, Phrygian).

The distribution of these nodes among the classical and pop subgroups (omitting the ‘neither’ subgroup) are shown in Figure 5. The significance of these data will be discussed further below.

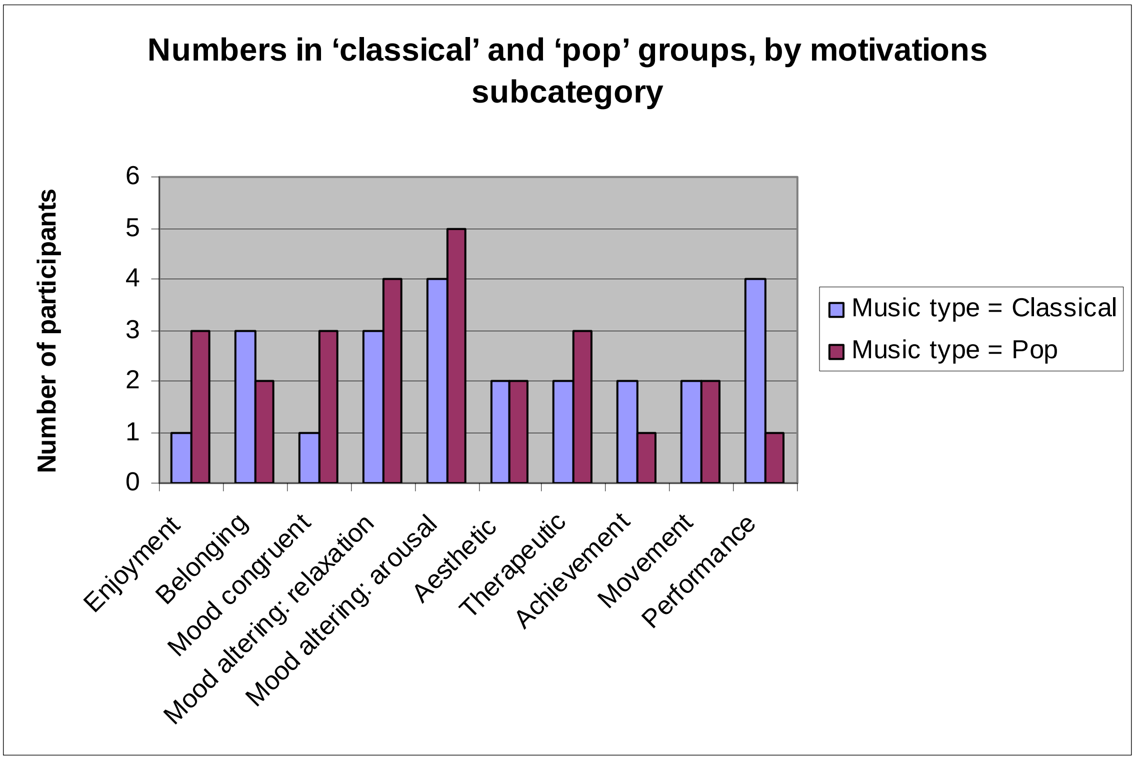


Figure 5: Numbers of participants in ‘classical’ and ‘pop’ groups, by motivations subcategory

3.4.3 What music did participants engage with: Characteristics

The node structure for this tree is given in Figure 6 below.

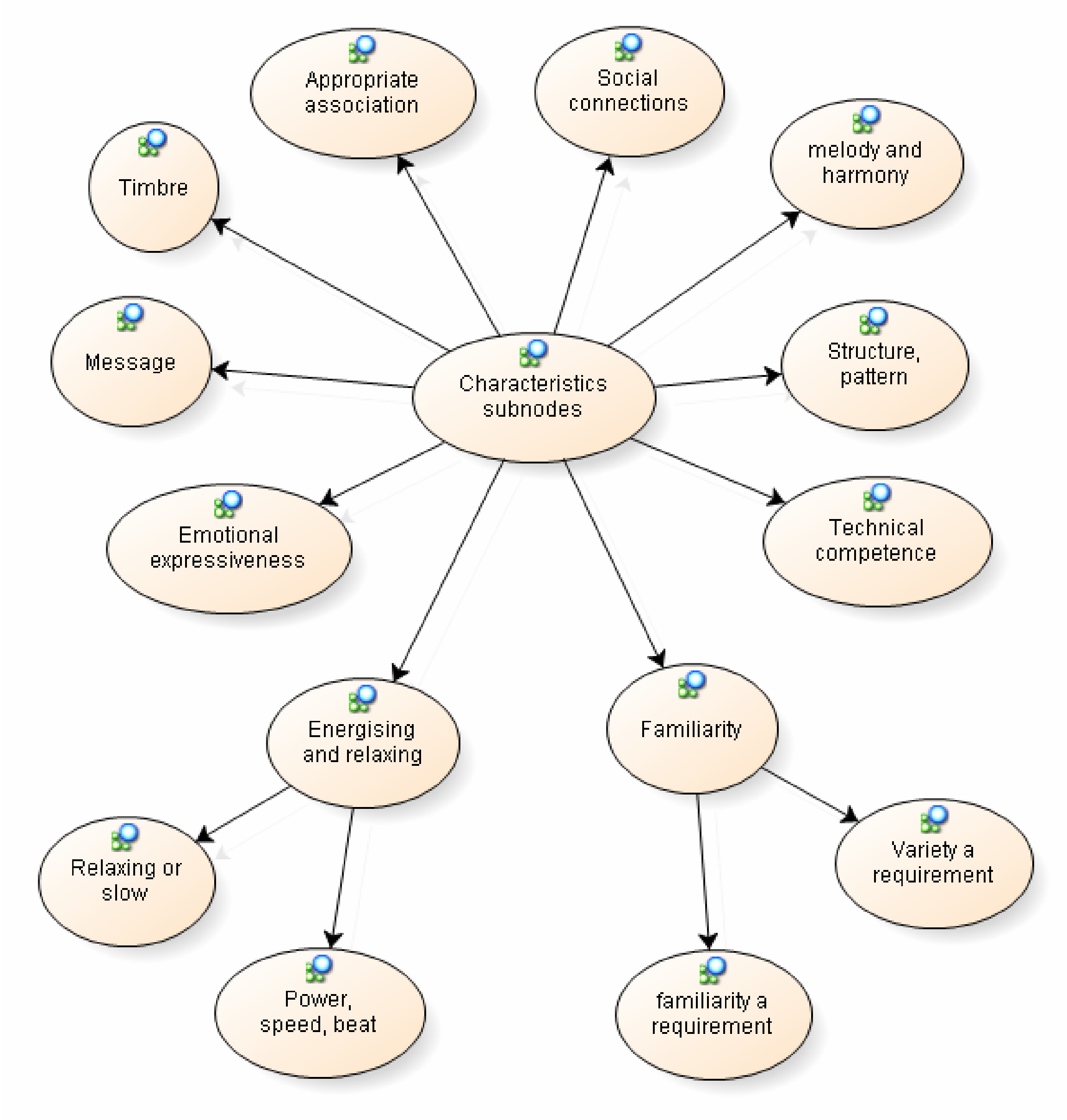


Figure 6: 'Characteristics' subnodes

Under the “characteristics” node, I created a subnode that included the features of music associated with its energising or calming effect. Ten out of twelve (%) participants described characteristics linked to this aspect, with participants enjoying rousing music; their descriptions included ‘go-ey’, ‘loud and lively’, ‘fast’, ‘exciting’, and they suggested it had ‘raw energy’, ‘power’, ‘impulse’, or ‘a lot of beat and rhythm’.

Five of these participants sometimes enjoyed slow quiet music to induce an effect of relaxation, when appropriate. However, several participants categorically disliked slow music, describing it as ‘dirgey’, ‘boring’ or ‘dragging on’.

Eight participants (67%) mentioned structure or pattern as one of the characteristics of the music they liked. Not all of these gave rise to entries in the “aesthetic” category of musical motivations mentioned above, because in some cases an appreciation of structural aspects was not the key reason for listening to it and therefore did not count as a “motivational” factor.

“Appropriate associations” were also important, being mentioned by eight people (67%). This category included cases where the significant factor was that the music was appropriate to the context, or because certain music induced pleasant thoughts. For one individual, music was only important in the context of film music, but seven others (58%), for whom music had a wider significance, also mentioned that film music was helpful in establishing atmosphere.

Emotional terms or associations, for example, “faster and happier”, “happy and forthright”, “plaintive”, “sad songs”, “slushy harmonies that instantly appeal”, and “cheerful or darker”, were mentioned by eight (67%) participants.

The importance of melody and harmony was mentioned by four participants, but melody was never mentioned on its own.

Dislike of change was mentioned by four participants, whilst for another four, variety was sometimes desirable (four made no mention of this aspect).

Social connections were mentioned by four participants, in the context of introducing them to music, for example: “I went to Australia when I was 20, and my friend in a record shop saved me a box of singles, when I heard some new groups”, and “Me and my mates Tony and Simon, we are very much into Phil Specter”.

The between group (classical and pop) distribution of these characteristics is shown in Figure 7.

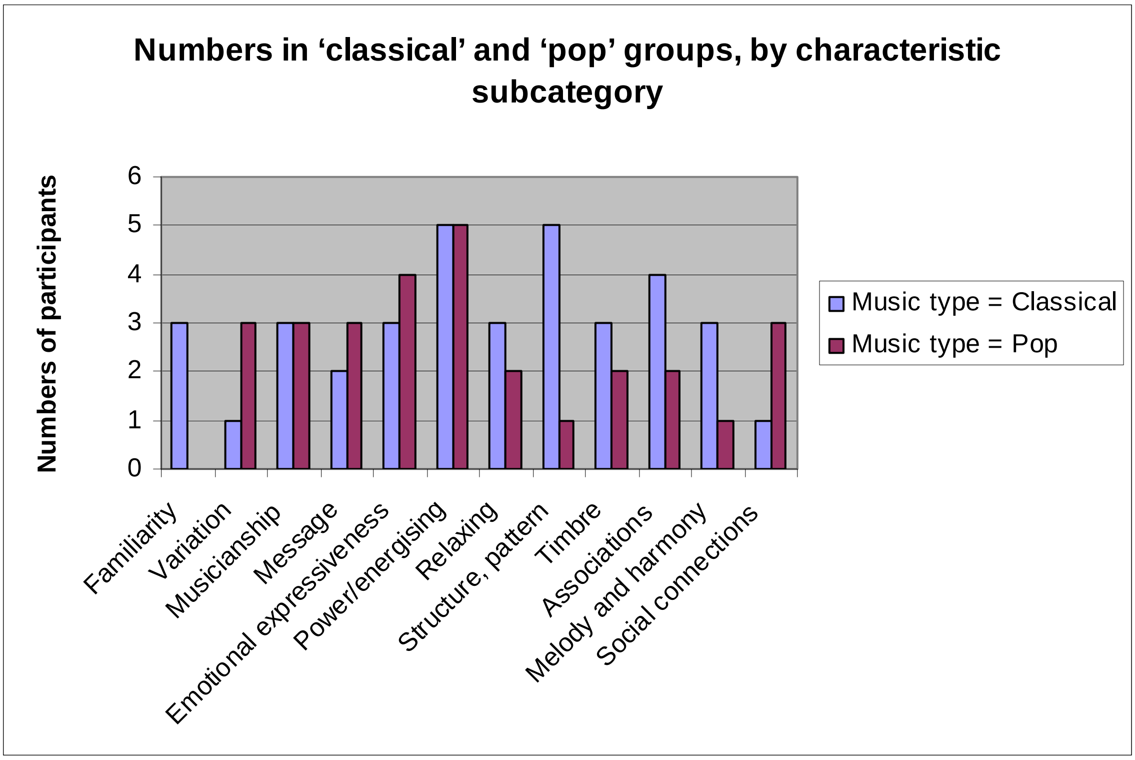


Figure 7: Numbers of participants in ‘classical’ and ‘pop’ groups, by characteristic subcategory.

3.5 DISCUSSION

Previous studies into music and autism have been limited to testing identifications between musical components and conventionally associated moods, and have provided no insights into the musical experiences of individuals with ASD or how these might differ from the TD population. In contrast, a growing body of research has studied how and why TD individuals listen to music, noting for example the importance of mood management, personal development and social affiliation (Batt-Rawden & DeNora, 2005; DeNora, 2002; Lamont, 2003; North, Hargreaves & Hargreaves, 2004). Whilst no direct comparison between ASD and TD samples can be made, the TD work provides a context in which music in autism can be considered, and compared with the data reported in the current study.

In contrast to previous experimental studies that provided total numbers of affective response categories, the present study adopted a methodology that was open-ended, as this was deemed to be the most effective way of eliciting information about personal experiences. Whilst this study does not have a matched control group, extensive prior qualitative research into the nature of the musical experience in non-autistic populations has been carried out, and provides a context for my analysis. Advantages of this design are that it is not limited by preconceptions or prior hypotheses, and conforms to standards of validity in qualitative research. Aside from that relating to music therapy, there has been little or no published work concerning the subjective experiences of music in individuals on the autism spectrum.

3.5.1 Motivations

3.5.1.1 Mood change

The fact that five participants explicitly denied using music to induce valency changes is consistent with statements by eight (67%) participants that they had difficulties in verbalising their own emotional states or understanding those of others, for example: “I find it difficult to talk about basic emotions because I don't think in terms of those emotions”, or “I don’t know a lot about emotions, or know whether that was the emotion I was intended to feel [when watching a film]”, or “Emotion is a vague area for me. I really can't quantify it, and I find it impossible to say what an emotion is. I can understand the basics, a smile or a tear, and apart from that it’s too complicated for me”.

This finding is in line with studies by Berthoz and Hill (2005) and Hill, Berthoz and Frith (2004) (but see also Fitzgerald & Bellgrove, 2006; Hill and Berthoz, 2006), concerning the prevalence of alexithymia in autism. Alexithymia, or being “without words for emotions,” implicates a cluster of cognitive and affective components in a specific type of emotional dysregulation. Central to alexithymia are difficulties in identifying and describing feelings and difficulties in distinguishing feelings from the bodily sensations of emotional arousal (Hill, Berthoz, & Frith, 2004).

Alexithymia is a common condition in ASD, possibly affecting as many as 85% of individuals (Hill, Berthoz, & Frith, 2004). Two variants of alexithymia have been identified, which have been labelled with admirable parsimony (if perhaps some lack of imagination) as types I and II (Bailey & Henry, 2007). Individuals with type I fail to experience or describe emotions, whereas those with type II show a normal or high degree of conscious *awareness* of emotions but a reduced capacity to cognitively *appraise* them. In other words, the person with type II alexithymia may experience a relatively normal range of emotions but will have a deficit in the ability to label or identify them or discriminate between their own internal emotional states. This latter type of alexithymia is the one most commonly observed in ASD (Berthoz & Hill, 2005).

In other words, the kind of alexithymia commonly found in autism involves a preserved ability to experience emotional arousal and to be conscious of such arousal in oneself, but a deficit, relative to controls, in one's ability to verbalize one's emotional experiences and to analyze one's own emotional states and reactions. The sample in the present study did indeed show a normal to high degree of conscious awareness of the emotional arousal induced by music, though they were generally limited to describing the effect in terms of excitement on some occasions or calmness and relief from tension on others, suggesting that they had precisely the difficulties in verbalizing and analyzing emotional experiences and poor insight into these experiences reported by Hill et al (2004) and Berthoz and Frith (2005). This reinforces the need to draw a distinction in ASD between the experience of emotions and the ability to analyze and to report on them, especially where emotional reactions to music are concerned.

The most interesting finding among the motivations subnodes, arose from the analysis of the largest subnode of the "motivations" category, involving the use of music to achieve changes in mood. This showed descriptions of desired states clustering around the two regions of exhilaration and calmness. Four participants described the state of tension as undesirable, but one which could be relieved by the calming effects of music. Participants were not explicitly asked whether they sought calmness as a relief from tension, but the fact that four nevertheless spontaneously mentioned the need to reduce tension when speaking about the need for relaxation, implies the existence of a mood axis or dimension, ranging from calmness to tension, among the participants. The arousing or exciting effect of music was not clearly identified as having an opposite pole, though one person did describe feeling "emotionally dead a lot of the time".

3.5.1.2 Therapeutic

The size of this category (five individuals, see Figure 5) shows that music does indeed have a healing value for some adults with ASD. This is consistent with a study by Batt-Rawden and DeNora (2005), which claimed that a number of TD individuals suffering from chronic illness had also gained therapeutic benefits from music.

3.5.1.3 Belonging

5 out of 12 (42%) of participants found benefits from this aspect of music. Social deficits are characteristic of ASD, and it would have been expected, on this basis, that socially linked benefits of music would be correspondingly reduced in line with the reduction in social skills in the ASD group, and that the "belonging" benefits might even be absent in the group. 42% can therefore be regarded as a comparatively large proportion in this connection. It appears that for this sample at least, music fulfilled a role in reducing the loneliness characteristic of ASD. It illustrates the fallacy, implicit in the quotations from Huron, Levitin and Peretz cited above, in assuming that the social deficits found in autism necessarily involve a lack of interest in feeling part of the wider community. While it is true that some individuals with autism show no desire to socialise and are quite content with their own company (and the same could be said for many people with no tendency towards autism), it does not logically follow that lack of aptitude for an activity implies lack of interest in it. It may be that the lack of social success often experienced by people on the autism spectrum does not deaden, but rather sharpens, their need to be close to other people. If trying to make friends on a human level so often involves being hurt and rebuffed, because a person with autism is unable to respond appropriately in the delicate ritual of social interactions that is the means by which two people build up trust and esteem for one another, this is all the more reason for engaging in an activity which provides some of the benefits of friendship without the anxieties and failures associated with dealing with the complexities of human

relationships. The evidence from this study shows that performing music, buying music and listening to music can provide precisely these benefits.

The first two of the quotations cited in the Results section under the heading of ‘belonging’, illustrate the feeling that association with a musical experience made the listener feel ‘part of something’. The third quotation under this heading added the interesting detail that the individual concerned was aware of belonging to a generation for which certain types of music helped to define their identity. Such feelings are a commonplace of musical experience in the wider population, but it is perhaps surprising to see them expressed so unambiguously within this group of individuals with ASD.

It might be expected that there would be an overlap between the ‘belonging’ set in the motivations category and the ‘social connections’ group within the characteristics category, but only one individual was in both sets. It may be that the need to feel ‘part of something’ is less for those with friends with whom they can enjoy and discuss music directly.

3.5.1.4 Performance

Three out of the five (%) individuals in this category played music in some public setting (e.g., in church), and enjoyed the act of making music rather than the fact that it was a public event. The first of the two quotations under the heading ‘performance’ in the Results section, from a person who, incidentally, described herself to us as being completely unmusical, illustrates that pleasure in making music could be quite independent of reciprocal public involvement. The second quotation, by contrast, shows the dominant role of audience involvement in the strong desire of one individual to perform publicly.

3.5.2 Characteristics

The salience of “power” as a desirable characteristic of music is consistent with the principal use made of music, to change mood, of which inducing a feeling of excitement or exhilaration is prominent. Powerful music seems likely to be an effective stimulant of such states. The dynamics of the music appear to be important in providing “power”: speed, beat and rhythm were specifically mentioned here. The importance of “structure” is consistent with the salience of the “aesthetic” use of music, though this admittedly is partly a consequence of the assumptions I made in defining the “aesthetic” category, namely that enjoyment of structure in music counts as an aesthetic reaction to it. The size of the “social” category shows that this aspect cannot be ignored in considering music and autism. The importance of “timbre” was unexpected, but is consistent with reports of enhanced perception in autism (Mottron & Burack, 2001) and merits further investigation.

It may be of interest that emotional terms or associations were used in describing music by eight participants, despite their claiming, elsewhere in the interviews, that they did not understand these terms when they concerned relations with other people.

Whilst the importance of melody and harmony was mentioned by four participants, melody was never mentioned on its own, suggesting that this musical component has a low profile in the musical vocabulary of this group of participants.

Dislike of change is often characteristic of autism. Indeed, this feature is such a common one in the condition that “need for the preservation of sameness” is sometimes regarded in its own right as a core dysfunction of autism (for example, in Deutsch et al., 2010). On this basis it might be expected that familiarity might be a requirement for favourite musical items, and it would not be unreasonable to expect it to feature in responses from 100% of participants. However, it was mentioned by only four (33%) participants in this study. In the circumstances, this must be regarded as a rather lower proportion than would be expected.

3.5.3 Profile differences between pop and classical groups

The differences in the profiles of the ‘pop’ and ‘classical’ groups for the overall category of motivations suggest that, as might be predicted, the more formal and perhaps musically deeper engagement of the classical group is reflected in an emphasis on the enjoyment of performance and the more intellectual/aesthetic side of music, whilst the pop group benefits somewhat more from the purely emotional charge of the music, with higher numbers represented in “enjoyment”, “mood congruent”, “mood altering”, and “therapeutic”. The greater emphasis on musical qualities such as structure, melody and timbre in the ‘classical’ compared with the ‘pop’ group might have been expected, given the higher overall level of formal musical training of the classical cohort. The greater importance of “social connections” for the ‘pop’ group, referring to the fact that friends had introduced them to new types of music, might reflect the more fan-based nature of this musical category. The importance of familiarity in the ‘classical’ group, and of variety in the ‘pop’ group, was unforeseen, but might be explicable on the basis that individuals whose musical tastes formed relatively late in development may have been more prepared to take risks with their musical choices as compared with those whose listening patterns were formed as children. Alternatively, it may reflect differences in levels of complexity or other variables in the music itself.

3.5.4 Reliability, Validity and Objectivity

These constructs are of limited applicability to qualitative research, though Spencer, Ritchie, Lewis, and Dillon (2003), especially on pp. 59-69, discuss ways in which they can be defined in qualitative studies. However, to those familiar with quantitative research, for whom reliability etc have specific technical definitions, it might be more informative to discuss the issues in terms such as dependability, credibility and confirmability.

In the formulation of Lincoln and Guba (1985), who first defined these terms, dependability means showing that the findings are consistent and could be replicated, credibility refers to providing confidence in the truth of the findings, and confirmability involves showing that the findings of a study are shaped by the respondents and not researcher bias, motivation, or interest.

The preceding analysis of results shows that in most respects, participants responded similarly to well established findings in the typical population. Where differences were observed, particularly in types of descriptive words used, these appeared to show a consistent trend towards adoption of a certain kind of expression, of an inward-directed rather than outward-directed nature, suggesting that dependability is not an issue in this study.

Credibility, to the extent that it is possible to have confidence in these findings, is supported by the fact that the observed differences are such as might be expected in the light of previous findings that autism tends to involve a more low-level perceptual processing style than with typical populations.

The claim of confirmability rests mainly on the qualitative method adopted in this study, Grounded Theory. It is a fundamental principle of this approach that the findings are based on the transcripts of interviews with the participants, in which "steering" of participants is kept to the minimum consistent with covering the topics included in the research question. The formation of nodes in NVivo arises as far as is possible from the raw material in participants' responses, and researcher bias and motivation have little scope for interfering with this process.

The formulation of the semi-structured questionnaire, it is true, involved elements of subjective choice. However, it is arguable that the precise formulation of the questionnaire was not critical to the outcome. Pidgeon and Henwood (p. 258) specifically warn against "becoming constrained by pre-formulated questions rather than adopting a more open-ended conversational style", and Spencer et al. include "listening rather than talking" (p. 59) as a prime consideration in what they term 'validity'. In the present study, the supplementary questions were available as prompts if the conversation flagged, but in most cases I found that the rather formal process of completing the early musical experience questionnaire broke the ice, and led on naturally to the discussion of how the participant's musical interests had evolved since childhood or adolescence.

This approach, via the participant's personal history, turned out to be highly productive. In many cases, participants would talk for ten or twenty minutes or even longer on their personal musical histories, making it unnecessary to do more than add the occasional prompt to explore further aspects of their experiences and restore the flow of the participant's conversation. The contamination of the results by the experimenter's preconceived ideas, which might be seen as a major threat to credibility in exploratory qualitative studies, was thus kept to a minimum. In addition, it is argued that the author's very lack of prior experience and knowledge of the area (a lack, which as has been shown above, is widely shared in the psychology community), made it even less likely that there would be a temptation to imprint preconceived ideas onto the data.

The fact that the autism group was a convenience sample limits the extent to which the results can be generalized to the total population of high-functioning adults with ASD, though as Accordino et al. comment with regard to analyses of music therapy: "while the research on more than one client may lack the detail of a case study, it is much more generalizable than a case report that only notes changes in a single participant".

Nevertheless, it should be acknowledged that confirmability is one area of limitation in this study. There was a degree of what qualitative theoreticians term "member checking" during the interviews, in the sense of summarising or restating assertions to ensure that participants had not been misunderstood. However, given that the main findings were obtained after the conclusion of the interviews, there was no opportunity to check these systematically with participants and thus no ability to triangulate the results. However, some of the hypotheses suggested by this study were developed and tested in a further set of experiments, described below in chapter 5.

3.5.5 TD populations: similarities and differences

My results showed that music was used by the sample in several ways that were similar to those reported in the TD literature, such as mood change, self-management for depression and social affiliation. However, one striking difference was found. TD individuals describe their mood changes in response to music as lying along two axes, one single axis of valency (happy/sad), and another single axis of 'arousal' (Bigand, Vieillard, Madurell, Marozeau, & Dacquet 2005). But the sample showed almost no use of valency terms, and instead used descriptors of states lying along two dimensions of arousal, with calmness/tension as opposite poles of one axis, and excitement or exhilaration as the desired state on the other axis.

This suggests similarities with a model described by Thayer (1978), which, however, expands the traditional concept of arousal from one to two dimensions rather than describing mood space. He identified dimensions of vigour/tiredness, and of tension/placidity. The latter clearly maps onto the tension/calmness dimension of these participants. The vigour/tiredness axis was not observed in its entirety here, though it is plausible that the node characterized by descriptors such as 'excitement' and 'exhilaration' is identical to the 'vigour' pole of Thayer's vigour/tiredness axis.

The fact that the ASD group consistently used terms which described internal states of arousal (calm, tense, exhilarated etc) to describe the subjective impact of music, rather than terms implying valence (happy, sad etc), is consistent with findings by Capps, Yirmiya and Sigman (1992) and Bormann-Kischkel, Vilsmeier and Baude (1995), that the facial emotions for which their autism group showed specific recognition impairments were those (such as embarrassment or surprise) with an external locus of control and which required knowledge of social scripts, rather than those which were mere signs of inner states. It is also consistent with the comment in Bowler (2007, p. 246) that "individuals with ASD seem to engage in less top-down processing when making perceptual judgements, that is to say, their reactions to the world are based on information that is closer to the properties of the incoming stimulus". Attributing emotional characteristics to music would seem to involve more top-down processing than describing changes in internal arousal.

On the one hand, the absence of a comparison group in this study limits the degree to which findings in this sample can be compared to typically developing populations. These findings must therefore remain somewhat tentative until they are replicated with a larger sample and a comparison group. It would also be of interest if a study of childhood descriptions of musical experience were carried out to enable a comparison between typically developing and ASD individuals at an earlier developmental stage. Notwithstanding these limitations, the current study still stands to date, to the author's knowledge, as the only in-depth qualitative study into the nature of the musical experiences of high-functioning adults on the autism spectrum that has been conducted on any scale larger than that of individual case studies, and that has been published in a peer-reviewed psychology journal.

3.6 CONCLUSIONS

I have shown that individuals with ASD can respond profoundly to music, and show considerable understanding both of the music and its effects on them. Whilst the group generally found it difficult to verbalise their emotional responses to music, other than in terms of internal arousal states, they clearly experienced a wide range of benefits from it. The participants in this sample, which was not selected for musical experience, exposure or knowledge, actively and consciously used music in their daily lives.

The two outliers in the subgroup whose engagement with music was minimal, seem to have levels of musical understanding which are in this respect relatively restricted: they may be characteristic of a subtype, of which Temple Grandin, who reports “not getting” music (Sacks, 1996), is an example. Without extending this study to a larger and more representative sample, it is impossible to say with certainty how large a proportion of the high-functioning ASD population may fail to engage with music, but on the basis of the current sample, it is plausible that it would apply only to a minority.

The majority of the sample found music of value in achieving improvements in mood, as well as for improving personal and social integration. This suggests that there are likely to be extensive practical benefits if caregivers or clinicians working with high functioning adults with ASD encourage them to foster their musical interests. I will return to a consideration of this in Chapter 6.

The next two chapters, 4 and 5, will deal with the design and execution of a series of experiments intended to follow up the principal findings of this chapter by developing and testing quantitative hypotheses by using a larger autism group and comparing the results with an age and verbal IQ-matched control group from the typical population.

Chapter 4. Piloting the quantitative measurement of reactions to music in autism.

Polonius: What do you read, my lord?
Hamlet: Words, words, words.

William Shakespeare (*Hamlet*)

4.1 SUMMARY

Chapter 3 described the process of exploring the subjective experience of music in adults with high-functioning autism. The value of qualitative methods in generating hypotheses for further testing is often enhanced by the subsequent use of quantitative techniques to test those hypotheses, using standard experimental designs and statistical methods of analysis. The next logical step for the current programme of research was, therefore, to formulate plausible hypotheses on the basis of the qualitative data collected and described in Chapter 3, and to assemble experimental stimuli which could validly be held to represent the constructs mentioned in the hypotheses. This chapter will outline the considerations which lay behind the selection of the stimuli used for these experiments; details of the procedures used to present these stimuli to participants, and certain other details such as the calculation of the physiological effects of music, will be left to the following chapter.

4.2 DEVELOPMENT OF HYPOTHESES

As already noted in chapter 3, the overall approach of our autism sample to music, and the ways in which they used music in their everyday lives, was strikingly similar to the results reported in other published research into the typical or non-autistic population. Although a useful purpose may have been served by pointing out these similarities in the context of a qualitative study, there are problems with developing this theme into a quantitative investigation. The most obvious of these is that in the standard form of the null hypothesis significance test, clear conclusions can only be reached in quantitative group comparison studies when a significant *difference* is identified, and statistical tests can never prove or even provide a high degree of probability for concluding that *no* difference exists. It therefore appeared that the most useful way to take this work forward into the quantitative domain would be to focus on the one area where a difference was apparent, namely in the use of somewhat limited and unsophisticated language by the autism sample to describe their emotional reactions to music.

Any examination of the use of language to describe emotional experiences in autism needs to take account of the prevalence of alexithymia in this condition. As was briefly discussed in chapter 3, the type of alexithymia generally found in autism involves a reduced ability to translate internal emotional experiences into precise, explicit verbal form. This suggests that in looking for differences between autism and control groups in terms of their reactions to music, one should be looking for the sort of differences which might be explicable in terms of the underlying alexithymia, which one would expect to be present in the autism group. In the musical context, one might expect that alexithymia would manifest itself in terms of impoverished verbal emotional responses to music. This impoverishment might be evident both in the form of a reduced and less explicit vocabulary in describing these responses, and in terms of the use of a smaller number of words in response to music. In other words, one would expect a person with autism not only to use fewer words to describe their emotional reactions to music, but also that these words would be drawn from a more restricted range of words, compared with non-autistic controls.

The results of the study reported in chapter 3 support the second of these expectations, in the sense that the words spontaneously used by participants to describe their reactions to music tended to be limited to words describing internal states of arousal, rather than externally focused emotion words. However, the fact that this was a qualitative study based on Grounded Theory, and the exploratory nature of the methodology, meant that it was not possible to formulate testable hypotheses until after the conclusion of data-gathering. It was possible therefore that the observed tendency towards impoverished emotional language in the autism sample might have been an artefact of the method of investigation, rather than a real point of difference between the general population.

However, the existence of this result, however ephemeral it might turn out to be, suggested that it would be useful, as part of the quantitative follow-up to this study, to investigate whether the profiles of word-choice did indeed differ in some systematic way between the two groups. More specifically, the first hypothesis to be tested was that the autism group would have a greater tendency to choose *internally focused* words to describe their emotional reactions to music, compared with a typical control group.

The second hypothesis was that not only the type, but also the *number* of words used to express emotional reactions to music would be greater in the autism group compared with controls.

However, in using musical stimuli to generate emotions in listeners, and in then interpreting the listeners' verbal reports (in whatever form) in terms of possible alexithymia, one is making an important additional assumption, which it is essential to test and verify. This is the assumption that if individuals from groups sampled from two distinct populations are exposed to the same musical items, they will, on average, experience the same or a similar range of internal emotions. If, in our case, it is discovered that the musical stimuli produce an impoverished set of responses from an autism group, it is necessary to be able to rebut the alternative explanation that the difference is due, not to alexithymia, but simply to a specific lack of emotional responsiveness *to music* on the part of the participants with autism. This precaution is even more imperative in view of the previous speculations cited above, that people with autism will be specifically impaired in their emotional responses to music. Despite the fact that the previous study found no real evidence for this, it remains a possibility which must be allowed for if any firm conclusions are to be reached from the quantitative study currently under consideration.

The possibility of confounding alexithymia with lack of emotional responsiveness to music suggests the use of some non-verbal measure of emotional responsiveness, which can measure emotional arousal independently of the verbal report measure which is bound to be affected by alexithymia. There are a number of alternative ways of measuring emotional responsiveness non-verbally, mainly involving physiological measures such as monitoring responses in the facial muscles, heart rate, and skin conductance (Bradley and Lang, 2000). Out of these alternatives, measurement of skin conductance, or galvanic skin response (GSR) as an index of physiological arousal, is a well-established method for checking levels of emotional arousal independently of verbal or behavioural responses, and is comparatively simple and non-invasive.

Another possible complication arose out of the tendency of individuals with autism to suffer from hyperacusis, or extreme sensitivity to sound. One study has found 18% of a sample of children with autism to suffer from hyperacusis, compared with none of the children in a matched control group (Rosenhall et al., 1999). This meant that a standard procedure in auditory experiments, of presenting the sound stimuli at precisely equal decibel levels to all participants, was unlikely to be ethically justifiable, since participants with hyperacusis were likely to find sound levels adjusted to be realistic for normal listeners to be painful or even distressing to them. However, if the musical stimuli had to be presented at different intensities so as to accommodate the comfort levels of participants, this might clearly introduce an unwanted confounding variable which could prevent conclusions being drawn as to group differences.

I therefore decided to introduce a separate, control stimulus, consisting of environmental sound effect noises of no musical value, to be presented at the same volume to any given individual as the music stimuli. This was intended to make it possible to partial out any individual variability due to sound sensitivity per se, and to measure not simply the effects of the music (which might involve reaction to the decibel value of the music rather than its aesthetic value), but also the effects of a simple auditory stimulus. Although the absolute value of the reaction to music might be subject to some unknown degree of interference from the factor of hypersensitivity, the difference between the reaction to music and the reaction to environmental noise stimuli would, hopefully, measure some real effect due to the music alone, that might be independent of varying sound levels and sensitivities.

Overall, then, the plan for the next phase of the study was to recruit age and gender-matched participants in two groups, an autism group and a control group, to expose them to a set of standard musical items, and a standard set of environmental noise items, while simultaneously measuring their physiological levels of emotional arousal, to derive their verbal descriptions of their emotional responses to the musical items, and to analyze whether the groups differed on either the profile or the absolute number of words used in the verbal descriptions after correcting for individual differences in the relative strength of the physiological arousal in response to the music compared with the environmental noises. Finally, the question would be addressed of whether any resulting group differences in either word profile or word frequency could be totally or partially explained in terms of the greater degree of alexithymia expected to characterise the autism group when compared with the control group.

4.3 OPERATIONALIZING THE HYPOTHESES

4.3.1 General considerations

In order to test hypotheses referring to the words used to express responses to music, the first step should be to specify the words and the music to be used. For a quantitative, controlled experiment, it is clearly advisable to use both a standard list of words, and a standard set of musical items, both of which should as far as possible be suitable for testing the hypotheses and representative of the constructs contained in them.

In operationalizing musical stimuli, one encounters difficulties specific to this particular area of investigation. At first sight, it might appear appropriate to use the musical stimulus in its simplest possible form, in order to make it easier to gain control over the complications and possible confounds that are introduced whenever complex stimuli are used. This suggests the use of purely synthetic stimuli such as pure sine waves. However, these raise such serious questions of ecological validity that they are seldom used in practice. On the other hand, if naturalistic musical stimuli are used, the selection of the stimuli involves a choice of a small number of items from among an almost infinite number of possibilities. These possibilities encompass music from such diverse styles as classical (itself subdivided into mediaeval, baroque, romantic, and modern), folk, soft rock, hard rock, jazz, soul, heavy metal, Latin, techno, dance and rap.

A further difficulty arises when considering the context in which the musical stimuli are presented in any experiment. Suppose for the sake of argument one is aiming to simulate the situation in which a participant might choose to listen to a favourite item of music for the purpose of mood regulation. The first point is that if a real life situation is being considered, he or she will be choosing the musical item. The choice of item may depend crucially on the type of mood regulation desired. The second point is that he, or she, will be deciding when and where they will be listening to the music. This choice, again, is likely to depend on the effect required, level of fatigue and other personal factors. In an experimental situation, on the other hand, the experimenter will probably wish to present a set of standard musical stimuli, so the decision as to musical item, and even musical style, is ceded to the experimenter. The experimenter will be asking the participant to listen to an item of music at a time set by the conditions of the experiment, not a time suitable for the purposes of the participant, so control over the time dimension is also ceded to the experimenter. Given the importance of locus of control to wellbeing and self-image (see eg Christopher et al., 2009), the reactions of the participant to music are likely to be deeply affected by the experimental context.

Now suppose the experimenter wishes to measure the emotional responses of the participant to the music, using verbal self-report. They have the alternatives of an open-ended format, in which the participant chooses their own descriptors (probably orally, but possibly in writing), or the typical "tick-box" format, in which the experimenter has set out a set of alternatives, with instructions which may allow the participant freedom to indicate just one, or an unlimited number of experienced emotions. The set of words will in most cases have been determined by a pre-existing theory of emotions, based on the theories of Ekman or Russell. The experimenter has additionally to take the decision whether they wish to measure the emotional response evoked by the music in the participant, or whether to focus on the emotional content of the music as perceived by the participant (glossing over the philosophical objections which might be raised over problems of defining "the emotional content of the music"), and finally to evolve a set of instructions which make his or her intentions perfectly clear to the participants.

It will be seen that the process of operationalizing the constructs involved in the measurement of musical emotions in such a way as to ensure that the outcomes have any degree of validity, is no easy matter. Not only is it hard to ensure that the variables are measuring what they purport to measure, but it is often problematic to compare the results of one study with those of another. One option which has been recommended in studies of emotional arousal is to have multiple measures, so that at least in cases where evoked emotion is in question, that construct can be validated by a kind of triangulation. This procedure is recommended in for example Bradley and Lang (2000). The difficulties with musical extracts have been harder to overcome. Some studies, eg Salimpoor et al. (2009), have finessed this by asking participants to bring along their own choice of musical items, namely those which caused them musical "chills" or moments of intense pleasure. While this escapes what might be called the over-standardization problem of providing musical items which the participant might not have chosen for themselves, it does limit the power of the design.

The multiple difficulties outlined above are potentially further complicated when, as here, the main focus is on an examination of group differences. To the variables of musical type and presentation, and measurement of emotional response, is added the further independent variable of group membership, with the corresponding requirement for some kind of control over possible confounding variables. These variables will almost certainly include age and gender, given the gender bias in autism and the tendency of most "control" participants in psychology experiments in UK universities to be women in their early 20's. Verbal IQ is another factor which has to be considered, given that if emotional responses are to be explicitly verbalized, the level of vocabulary of an individual might well affect their tendency to identify which of a list of words describes their emotional state. The latter type of difficulty, that of matching the autism and control groups, is at least one which is well understood and whose solution involves fairly standard procedures. This issue will be dealt with in chapter 5.

The problems with musical items and with measurement of emotions were not so simply dealt with. Given that the sample size in the autism group was likely to be limited by my access to high-functioning adults on the spectrum who were willing and able to participate and who would be a match for my control group on the measure of vocabulary, a design was needed which would produce a comparatively large effect size, to make up for the small sample size in the autism group. This suggested the requirement for some kind of rigidly standardized procedure in order to minimize the levels of unwanted error or "noise", so as to allow any "signal" indicative of group differences to be more clearly discernable. At the same time, this would make it difficult if not impossible to base an experiment on using the participants' personal favourite items of music, which was the solution to the over-standardization problem in the design of Salimpoor et al. (2009). The remainder of the current chapter describes the development of the experimental procedure and selection of musical stimuli. Full details of the study that this led to will be described in Chapter 5.

4.3.2 Compilation of the musical items

In order to allow for the difficulties inherent in the study of music and emotions as well as sample size and effect sizes outlined above, it was decided to use a set of standardized extracts.

Selection of the ‘fixed’ musical excerpts for the study was aided by the work of Eve-Marie Quintin, in the sound laboratory at McGill University, Montreal (Quintin et al., 2009). Quintin had piloted a total of 28 music clips to target the emotions of happiness, sadness, fearfulness and peacefulness. The duration of the clips ranged from 30 to 50 seconds. The extracts had been piloted with 20 undergraduate and graduate students, and only those 20 clips were retained for which the highest level of inter-rater agreement had been obtained. The clips were purely instrumental and non-vocal, thus eliminating the possible confounding effects of human voices and, in particular, lyrics, which might convey emotional messages independently of the content of the music itself.

Out of these clips, only those that had been identified by her pilot participants as being, in the terms used by Quintin, “happy”, “sad” and “scary” were chosen. Four items were taken from each of these three categories, being those identified in her pilot testing as being the ones whose emotional connotations had most frequently been identified correctly (Quintin, personal communication). These conform to the three basic emotions which are known to be universally recognised in music (Fritz et al., 2009), and are all among the six basic emotions described in Ekman and Friesen (1976), whereas “peacefulness” is not among the list of basic emotions, nor has it yet been described as recognized by similar features in music cross-culturally. The items all consisted of fairly standard items comprising a range from film music (well represented by Hitchcock for the “scary” items) to a Bach prelude. There were no items falling clearly into the sort of pop categories which would be likely to appeal strongly to younger participants, however. The length of the items selected was in all cases within plus or minus 3 seconds of 30 seconds. The loudness of the stimuli had been adjusted by Quintin on the basis of loudness matching judgements in a separate pilot study in order to create perceptually equal stimuli. A list of the musical items from which these extracts were chosen is given at Appendix 4.1.

It should be noted that the emotions “targeted” by Quintin were emotions perceived by listeners in the music, rather than emotions evoked by the music. The question of whether the items evoked the same emotions as they expressed, was one which was resolved later in the piloting process, when verbal descriptions of the effects of the music were elicited from participants (further details of this process are given below). The outcome was that the emotions described by participants as induced in them by the music, when the words used were categorised into the three possible alternatives of Sad, Happy, and Fearful, corresponded exactly to the categories described by Quintin et al.

The environmental sound items were selected from CDs published by the BBC sound library. Six items were chosen, on the basis that they were free of any sudden or potentially unpleasant sounds. The selected items represented rainfall, running water, the sound of a cross-channel ferry, a quiet city street, a bicycle on gravel and wave sounds. Each selection was 30 seconds long, matching the length of the music items.

4.3.3 Compilation of the word lists

My plan at this stage was to develop a list of emotion descriptors that would be used for testing participant reactions, which would achieve two aims. The first of these was to test whether the words chosen by the autism group tended towards the use of internally directed emotions more expressive of states of arousal, as the exploratory study had suggested. The second was to see whether the absolute number of words ticked in the boxes showed any group difference, as would be expected if the autism group showed a generalization of its trend to alexithymia by exhibiting also a reduced ability to name the emotions evoked by music. It was of course possible that both trends would be observed.

In order to satisfy both aims, I began by compiling a list of the words already used by the autism participants in the exploratory study to describe their affective reactions to music (see Chapter 3). These comprised words describing arousal states, such as calmness or exhilaration, and numbered fourteen in total. At this point, I needed to pilot a similar list of typical emotion response words from a “control” group of people without autism. One way of doing this might have been to conduct an interview with the pilot control group, and examining their musical interests and collecting the words which they spontaneously used to describe their reactions to music, as I had done with the previous autism sample. However, this would have been time-consuming, and given that the work was no more than one part of a piloting process for another experiment, it would have been difficult to justify the expenditure of time. Moreover, the more important hypothesis for test in the following study was that of determining whether control participants would tend to use more words than the autism group to describe their affective reactions to a standardised set of musical items. This part of the study did not require that the set of words used should have been obtained under identical conditions from the autism and control samples, merely that both groups should be tested with the identical set of words from which to select those describing their reactions.

The task of obtaining a "control" list of words to describe non-autistic responses, was undertaken by piloting the 12 musical items with a group of fifteen controls and asking them to list, for each separate item, up to five descriptors (ie words or short phrases) that described the participant's subjective feelings, thoughts, emotions or moods that the music aroused in them, even if only fleetingly, while listening to each item. I allowed participants complete freedom to describe thoughts, images or ideas, on the grounds that music often does result in the production of images rather than emotions.

When the responses for each item were compiled together, this produced a long list of adjectives, and some images or other associations, for each item of music. I eliminated words that too specifically referred to specific imagery associated with non-emotional aspects of the music (eg "Sesame Street", or "French Court"). The 14 most frequently used among the remaining words were then extracted. By this stage, I therefore had 14 words previously used by the autism sample in the study referred to in chapter 3, used spontaneously to describe their general reactions to music, and 14 words used by a non-autistic set of people to describe their specific reactions to twelve standard items of music (see Appendix 4.2 for the lists). There was a group of 14 words related mainly to internal arousal levels, and 14 words which tended to include more sophisticated emotional descriptors such as "wistful" and "dramatic"; call these group A words and group B words respectively. Given that the two groups of words had not been derived under identical circumstances, this would limit the conclusions that could be drawn from any group differences that might emerge in the profiles of word preferences. However, it would be reasonable to examine the profiles of word preference between the groups, in order to see whether there was any statistically significant difference in the proportion of group A words vs. group B words when the autism group was compared with the control group. Although groups A and B are not strictly comparable, it would still be of interest to see whether they differed in some way that was reflected in differences in word preference between the groups of participants.

4.4 TESTING THE VALIDITY OF EMOTION TERMS

Thus far, the process of obtaining emotion terms to describe items of music had all been one-way, comprising what could be seen in mathematical terms as a function, or mapping, from the set of twelve musical items into the set of possible descriptive words. The proposed experiment involved essentially repeating the situation that had already been piloted, namely of implementing a mapping by participants from musical items into word lists, with the important difference that in the actual experiment, as opposed to the piloting of it, the list of available word choices would be cut down from (potentially) thousands to precisely 28.

However, to anyone familiar with the mathematical idea of a function, it might occur to them to see whether the function was in any sense invertible, so that participants might not only reliably describe an item as having certain characteristics, but also be able to identify a particular musical item from other items on the basis of these very characteristics.

This would not only be an interesting question to answer in the mathematical sense (of discovering whether the function was invertible), but it might also help to provide evidence for the reality of the emotions evoked by the music. For if the process of describing the feelings supposedly evoked by music were in some sense an artefact of the experimental situation, or a demand characteristic, with participants cheerfully confabulating emotions which they do not really feel, one would expect it to be very difficult for them to work the confabulation process in reverse. By analogy, it would be a simple matter, given a set of random ink blots, to give descriptions of supposed images that one saw in them (whether or not these descriptions represented anything more than a desire to please the experimenter), in terms of, say, “a whale” or “a camel” or “a flower”. But if it turned out that participant B was reliably able to identify which ink blot had been described by participant A as “a flower”, and which had been likened to a whale, one would feel a much greater degree of confidence in the objective nature of the identifications.

Such considerations as these led to the incorporation of a sort of “backwards looking” aspect to the emotion naming experiment. It was decided to use only six out of the twelve items as probes to elicit the presence of evoked emotions, using the pre-determined list of 28 words and a tick-box arrangement. The remaining six were reserved, in effect, as tests of the validity of the whole emotion-naming exercise. For this purpose, instead of being presented with the tunes and being asked to identify appropriate words, participants would be presented with the words, and asked to identify the appropriate tunes. At this stage no decision had been made as to which six tunes to pick for this purpose, so the next stage was applied to all 12 musical items.

In practice, it was likely to be very difficult for participants to identify individual tunes, even in a set of only six tunes, on the basis of individual words. I therefore developed a set of around five descriptors per tune for each of the 12 items, comprising words which had been used by members of the pilot group for that particular tune. My supervisor, Dr Hill, chose her own set of five descriptors for each tune, and we separately tested whether another pilot group was able reliably to identify the tune corresponding to each “bundle” of words. For each bundle, the version (Dr Hill’s or my own) was selected that achieved the highest success rate for correct identification. The six tune/bundle combinations were chosen that achieved the highest rate of correct identification at this pilot stage, bearing in mind the need to maintain the balance between happy, sad, and scary items, with two of each being included in each of the two sets of six musical items. Overall, a rate of well over 50% accuracy was obtained overall in terms of correct identifications, which is well over the 16% that would be expected by chance alone (to anticipate the results of the main experiment, the average correct score over the total of 47 participants was 60.3%, a result whose probability on the null hypothesis of chance performance is so small, as to amount in practical terms to an impossibility).

The availability of “reverse” items not only allowed for a test of the validity of the emotion-naming procedure, but also gave scope for checking whether the autism and control groups might differ in their ability correctly to perform this reverse identification task. The instructions for this part of the main experiment would give the participants the best possible chance to get the correct correspondences, by making clear that the aim was not to describe their own reactions, but to guess how a previous group would have reacted to the items; in other words, it would be a kind of “Family Fortunes” situation. The list of “bundles” eventually chosen, and the six tunes to which they correspond, are given in Appendix 4.3 below.

4.5 DEVELOPMENT OF INSTRUCTIONS

The remaining part of the piloting operation involved developing a format for the test responses, and a set of instructions, which would be most easily comprehensible, and least susceptible to misunderstanding. This was done by “testing” a pilot group of individuals who had not taken part in the preparation of materials described above, and asking them to describe in their own words what they had understood the instructions to mean. This process ironed out a surprising number of problems, and the final form of the experiment resulted in a slimmed down and simplified procedure that seemed to be fairly universally understood. The form of words used is at Appendix 4.4.

At this point all materials were ready for the study to commence, details of which are covered in Chapter 5.

Chapter 5. Analysis of the comparative effects of alexithymia, musical experience and physiological responsiveness on levels of self-reported emotional response to music

I don't want to hurry it. That itself is a poisonous twentieth century attitude. When you want to hurry something, that means you no longer care about it and want to get on to other things.

Robert Pirsig (*Zen and the Art of Motorcycle Maintenance*)

5.1 SUMMARY

This chapter will present the data from the final study of the thesis. The principal aim of the work reported in this chapter was to compare the emotional reactions of individuals with and without autism to a series of standardised extracts of music, and to a control condition consisting of a series of loudness-matched environmental sounds. Emotional responses to the music were to be measured both physiologically, using the galvanic skin response (GSR), and verbally. It had been hypothesised that individuals with autism were likely to show a similar degree of *physiological* emotional arousal to music as controls (when differences in sensitivity to sound generally, as measured by the environmental sound responses, had been controlled for), but that alexithymia, the inability to put these emotions into words, would limit the extent to which they would report *verbally* on their emotional responses. The studies reported in this chapter support the prediction that basic, non-verbal responses to music in the autism group differ little from those in the control group. This conclusion is derived from two converging lines of evidence. Firstly, physiological responses to standard items of music are similar in the two groups. Secondly, group differences between the verbal emotional responses to music are substantially accounted for by group differences in alexithymia.

5.2 ALEXITHYMIA: DEFINITION AND MEASUREMENT

Before discussing details of the experiment, it will be necessary to discuss both the definitions of the two types of alexithymia under consideration, and the self-report measures that are used to measure them.

Alexithymia can be subdivided into two subsidiary concepts. The concise formulation given in Berthoz and Hill (2005) is as follows: "Type I alexithymia is characterised by a low degree of conscious awareness of emotional arousal and a low degree of emotion accompanying cognitions. Type II alexithymia is characterised by a normal or high degree of conscious awareness of emotional arousal together with a low degree of accompanying cognitions" (p. 296).

It will be seen that the two concepts have a common element: "a low degree of accompanying cognitions", in other words, a low score on the "cognitive dimension" of emotional experience, which might therefore be seen as the core diagnostic criterion for alexithymia. But they differ on "the degree of conscious awareness of emotional arousal", or what is – to anticipate – defined below as the "affective dimension" of the emotional experience. To the extent that "degree of conscious awareness of emotional arousal" is well defined, the two alexithymia types are therefore non-overlapping: the degree of conscious awareness of emotional arousal level of a particular individual must be either "low" or "normal or high", but cannot be both.

To paraphrase, type I alexithymia is associated with the phenomenon of individuals who do not experience strong emotions at any conscious level, and therefore, *a fortiori*, do not have the corresponding cognitive experiences. Type II is associated with people who have the emotions, and are aware of having them, but do not have the corresponding cognitions: in other words, they have feelings, but may not know, or be able to describe, what they are feeling or why.

It could conceivably be argued that even someone with type I alexithymia may not be totally emotionally unreactive, and that they might experience emotions at some subconscious level. However, for the purposes of this study, it should be made clear that the only aspects of the emotional experience that are considered, are those accessible to conscious awareness. It is assumed that physiological reactions (sweating palms, butterflies in the stomach, goose pimples etc) come into this category even if the person experiencing them is not aware of why the experiences are occurring.

Alexithymia is normally measured using two validated self-report questionnaires, the Toronto Alexithymia Scale (TAS-20: Bagby et al., 1994) and the Bermond Vorst Alexithymia Questionnaire (Vorst & Bermond, 2001). There are in fact three versions of this latter instrument, designated as the BVAQ-A and BVAQ-B with 20 items each, and the BVAQ-(A+B), with 40 items. In the present study I have followed Berthoz and Hill (2005) in using the BVAQ-B to the exclusion of the others, since it is believed to be somewhat superior to the BVAQ-A and to have the benefit of brevity over the BVAQ-(A+B). Although both the TAS-20 and the BVAQ-B can be scored to give a single overall value for “alexithymia”, they differ significantly in their detailed factor structure, and this difference is linked crucially with the distinction between type I and type II alexithymia.

At the time when the TAS-20 was created, no distinction was then made between type I and type II alexithymia. The subsequently developed BVAQ-B, however, was specifically designed to make such a distinction. In a nutshell, the TAS-20 measures what one would now call type II alexithymia, whereas the factor structure of the BVAQ-B enables us to measure type I and type II alexithymia separately. Not only does this give the BVAQ-B the edge over the TAS-20 in its coverage of the alexithymia construct, but there is also evidence for the technical superiority of the BVAQ-B over the TAS-20. For example, it has been found (Deborde et al., 2004) that the BVAQ-B is less sensitive to variations in participants’ emotional states than in the case of the TAS-20, so that the BVAQ-B is a better measure of alexithymia. To anticipate, this is entirely consistent with findings, reported later in this chapter, which conclude that the BVAQ-B, or more specifically one of its factors, is the best independent measure of the effects of alexithymia on the ability of individuals to identify the emotions induced in them by music.

The factor structure of the two questionnaires is as follows. The TAS-20 has three factors, labelled, respectively, as “difficulty in identifying feelings”, “difficulty in describing feelings” and “external [sic] oriented thinking”. The five BVAQ factors are "poor verbalizing", "poor fantasies", "poor insight", "poor emotional excitability" and "concrete thinking". Of these, factors 1, 3, and 5 together are said to represent the "cognitive dimensions" of alexithymia, and the remaining two – poor fantasies and poor emotional excitability – measure the "affective dimensions".

It is claimed that the affective dimensions together measure type I alexithymia, or, simply put, the ability to feel emotions. It is further claimed that the cognitive dimensions together measure the degree of type II alexithymia, and that they correspond factor by factor with the three factors in the TAS-20, which also, therefore, measures type II alexithymia. The detailed correspondences, established by Vorst and Bermond (and cited for example in Deborde et al., 2008) are as follows:

TAS F1 = BVAQ-B F3
TAS F2 = BVAQ-B F1
TAS F3 = BVAQ-B F5.

Given this correspondence, which – to anticipate again – this study broadly confirms, it is clear that the three TAS factors must between them measure the cognitive dimensions of alexithymia, i.e., the tendency to type II alexithymia, which are those comprised in the corresponding factors (1, 3 and 5) of the BVAQ-B. It is the presence of the additional two factors in the BVAQ-B, factors 2 and 4, which enables it to distinguish the presence (or absence) of type I alexithymia. In the absence of such factors, the TAS-20 simply measures the strength of the cognitive dimensions.

Therefore, if an individual were to score poorly on the TAS-20, it could be concluded that they exhibited a low degree of fluency in their emotional language. But this information would not throw any light on the question of whether this deficit was due to an absence of emotions – so that the poverty of language was due to the lack of anything for the language to describe – or whether it was due to an inability to put verbal labels onto emotional experiences which were real and present in the consciousness of the individual. It is precisely this distinction which the BVAQ-B is capable, it is claimed, of clarifying.

By exploiting this greater degree of discrimination, Berthoz and Hill (2005) showed that ASD was characterised by a significantly higher level of type II alexithymia compared with controls, as measured with the TAS-20 and with the cognitive measure (ie combined cognitive dimension factors) of the BVAQ-B, but that no significant difference was found in type I alexithymia, as measured by the combined affective dimension factors of the BVAQ-B. The key deficit in emotion recognition in ASD, therefore, appears to lie not in the existence of emotions, but in the ability to identify and name them. This distinction is important in what follows.

5.3 MEASURING THE DEGREE OF AUTISM USING THE AQ

The Autism-Spectrum Quotient (AQ) (Baron-Cohen et al., 2001) has been developed as a brief (50 item) instrument to measure the degree to which a normally intelligent adult possesses the characteristics associated with autism. As a one-dimensional scale, the use of the AQ requires us to accept the validity of placing individuals, on the basis of their AQ scores, somewhere along a linear measure of "autism": in other words, it presupposes a unidimensional "spectral theory" of autism.

Given that autism is defined in terms of a triad of impairments, this represents an assumption that requires proof, and indeed, for this and other reasons, the AQ is not recommended for use as a diagnostic tool, though subsequent research has confirmed its value at least for screening purposes (e.g., Sizoo et al., 2009; Woodbury-Smith et al., 2005). These studies have found that in practice, a conventional cut-off value for determining likely membership of the autism spectrum of 26 or more on the AQ has given results which correlate highly with the usual methods of professional diagnosis. The AQ has the advantages of being quick and easy to administer compared with the more clinically rigorous alternatives, and it was adopted in this study as a means of providing a greater level of information about the locations of participants on the autism spectrum than would be provided by the simple categories of "diagnosed with an autism spectrum disorder" or "control, not diagnosed with ASD". It also provided a check against gross error in terms of mistakenly including individuals with undiagnosed autism in the control group. In the event, given that the maximum AQ score encountered in the control group was 27, only just above the dividing line, and given that this score is in fact statistically within the normal range found by Baron-Cohen et al. (their sample had a mean of 16.4 and an s.d. of 6.3, so that the individual concerned was only 1.7 standard deviations above the mean), no control participants were excluded by this criterion. However, inclusion of the measure as a covariate did serve to sharpen some of the results, as reported below.

5.4 EXPERIMENTAL MODEL AND HYPOTHESIS

The following diagram, in Figure 8, shows the hypothesis to be tested, in graphical form. The "steps" refer to the transitions between one event (as described by the text within the ovals) and the next. It is these steps that are essentially the focus of the study reported here, and these will be referred to at various points below.

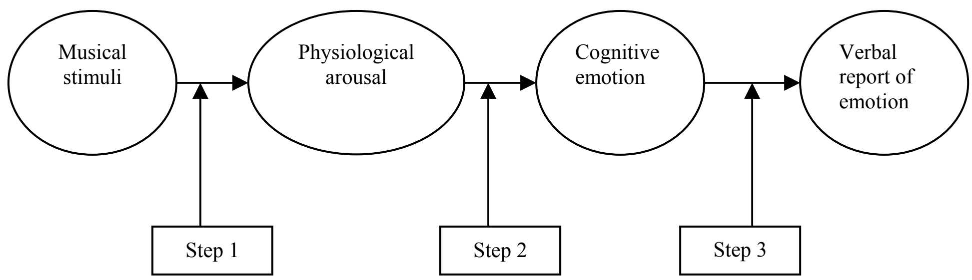


Figure 8: Schematic model for hypothesis under test

I predict that step 1 will be relatively unimpaired in autism, given that in the study described in chapter 3, the participants reported the usual benefits from mood change and feelings of well-being from listening to, or performing, music, and given also the apparent lack of type I alexithymia in their ASD sample reported by Berthoz and Hill (2005). These findings lead us to suspect that the basic ability consciously to experience emotions will be intact in autism, which requires that the most fundamental mechanism in emotional experience, that of physiological arousal, will also be in good working order. On the other hand, it is predicted that transmission of information at either step 2 or step 3 will be reduced by the effects of type II alexithymia, which will operate to reduce the ability to confabulate or verbally identify an emotion based on its physiological substrate. Note that according to this model, it is proposed that the effect of autism, if any, would be mediated by its association with alexithymia, and will not necessarily exert any independent influence. The effects of type I alexithymia, if any such effects were present, would manifest themselves at step 1, and those of type II alexithymia at steps 2 and/or 3.

It might be objected at this point that step 3 is in fact redundant, on the grounds that an emotion which cannot be described (ie verbally reported on, the end state of step 3) cannot have a cognitive component, and is merely physiological. Yet this is an assumption which is open to challenge, and which at this stage it is preferable not to make. It is surely possible, at least theoretically, that an emotion may be cognitively accessible, in the sense that one understands its cause and the action tendencies to which it gives rise, without one being able to give a precise verbal description of the feeling state to which it corresponds. For one very simple example, consider laughter. This is a response to which, presumably, an internal emotional state must correspond, and one which is occasionally intense. While it is true that laughter is not universal among humans (Voltaire reportedly never laughed, and according to some traditions, Christ never laughed), it is nevertheless a familiar enough subjective experience for most of us, and one which one should have no difficulty in identifying as the emotional state associated with mirth. Yet an accurate verbal description of the affective process which one undergoes when laughing is notoriously hard to establish. Hobbes' characterisation of laughter as arising from an internal feeling of "sudden glory" seems less than totally convincing, and the goal of pinning down its essence remains elusive to this day.

One would expect that a further factor affecting, in particular, step 3, would be the knowledge of vocabulary of a particular individual. Some evidence to support this expectation is provided by Heaton et al. (2008), which found that the ability to name the emotions expressed by music in a sample of children with autism and Down syndrome correlated significantly with level of language development. It was decided to control for this possible confound in the experimental design by testing the performance of the groups according to their performance on a receptive language test (BPVS: British Picture Vocabulary Scale, Dunn et al., 1997). The groups were not in fact matched precisely at the selection stage on their BPVS scores, but in the event it turned out that they did not differ significantly on this measure, and that the effect of this confound could therefore be disregarded.

A simple "tick-box" paradigm, as shown in the form given in Appendix 4.2, would yield an output corresponding to the right hand end of the process indicated above, i.e., it would measure the "verbal report of emotion" for participants. Standardizing the "musical stimuli" at the left hand end of the chain meant that any statistically significant differences between groups must be due to differences occurring somewhere between these two extreme points in the chain, assuming that this model of the process was correct. This means that the difference would be occurring in the transmission of information at steps 1, 2, 3, or any combination of these. Clearly this form of assessment, on its own, would not allow us to test this model in sufficient detail, for which one requires the ability to isolate the effects of autism (operating through the mediating effects of alexithymia) on each individual step in the process.

Taking separate measurements of the processes corresponding to the individual steps 1 to 3 would have two benefits. Firstly, it would serve to localise the effects of any group differences observed, and enable the differences to be described in some detail. Secondly, it would provide a cross-check on the validity of the model itself. In other words, triangulating the process through the use of multiple measurements would show whether the process was even plausibly described by the simple linear model given above; or whether the lack of any correlation between the results and what the model would predict, proved that some more complex hypothetical structure would be necessary. Either alternative appeared likely to yield results which would be equally interesting, if not perhaps equally gratifying. In the following section, I will describe how I tackled the issue of measuring performance at each step, and examining both the validity of the model and the group differences that it revealed.

5.5 EXPERIMENTAL DESIGN

I was guided in this area by the sort of considerations set out in Bradley and Lang, 2000. They state that emotions involve multiple responses, or an "emotional complex", which can be grouped into three broad categories: functional behavioural sequences, use of emotional language, and physiological changes. They go on to say that “although the emotional complex defies exhaustive measurement, it is imperative that researchers not limit themselves to the assessment of a single response (eg, subjective report). As a bare minimum, an experiment should include a sample measure from each major system: overt acts, language and physiology” (Bradley & Lang, 2000, p. 245). As will be seen below, this final experiment did indeed, so far as resources permitted, conform to this desideratum.

Step 1 corresponds to the degree of emotional responsiveness to the items of music, which represents the physiological element of Bradley and Lang’s desiderata in their comments cited above. It is traditional to measure this using the galvanic skin response or GSR (sometimes referred to as the skin conductance response or SCR), a measure of arousal of the sympathetic part of the autonomic nervous system (ANS). This has been regularly used to measure reactions to music, in, for example, Khalfa et al. (2002), though not, before the present study, in the context of investigations involving alexithymia. The rationale of the GSR method is that arousal of the sympathetic nervous system by any kind of emotional response leads to an involuntary activation of the sweat glands in the skin of the fingers, and a consequent reduction in the resistivity of the skin when a pair of electrodes are applied across two of the fingers of one hand. GSR measurements detect this reduction in resistivity, and consequently act as a proxy measure of emotional arousal.

GSR measurements appear to give robust responses to both positive (happy) and negative (fearful) responses to music (Khalfa et al., 2002). The analysis of the data is not always a simple matter, however. The most basic experimental setup uses a direct current source, which can be a small, nine volt battery. However, over the course of time, this can build up a polarisation potential across the electrodes, leading to a steady increase in apparent resistance and a steady apparent reduction in ANS arousal over time. A similar increase in skin resistance can of course be observed in participants who become accustomed to the experimental surroundings, which may at first have caused a degree of anxiety, which tends to disappear as habituation sets in.

In practice, the difficulty in dealing with these two confounds led to the adoption of a different approach to the data analysis of the physiological measurements. This process, which is explained in detail in appendix 5.1 below, involved calculating what was essentially a second derivative (in the calculus meaning of the term “derivative”), at each transition from silence to sound stimulus, and from sound stimulus to silence. It therefore measured the sharpness of the “elbow” in the GSR trace which was frequently present (and in many cases visually obvious from the trace on the computer screen of the GSR resistance readings) every time the stimulus changed from the presence, or absence, of the sound stimulus to the opposite condition: one might call it the “shock value” of the change in stimulus.

It was, in addition to the music items, necessary to provide a control condition involving the exposure of participants to environmental noise, to ensure that the physiological reaction observed to music was not simply due to exposure to sound at a certain decibel level, a reaction which would be of no real interest in determining the effect of music, as such, on emotional state. Whilst current definitions of music may be somewhat elastic, for the purposes of this study, I decided (perhaps somewhat arbitrarily) that such contributions to contemporary culture as 4' 33" by John Cage should not be included among the stimuli.

In order to measure the extent to which music had a physiological effect over and above any impact simply due to its decibel level, it was proposed to take the ratio of the music response to the environmental noise response, referred to below simply as “Ratio”, as being an operationalization of this construct with good face validity.

Further consideration/discussion of the experimental manipulations made to address the usual confounds is made later in this chapter.

5.6 METHOD

5.6.1 Materials

In order to provide the maximal information on the degree of alexithymia using conventional self-report measures, participants completed the TAS-20 and BVAQ-B alexithymia questionnaires. They also completed the AQ autism questionnaire. Details of these were given above.

5.6 2 Participants

A total of 23 high-functioning adults with diagnoses on the autism spectrum took part in the experiment. All diagnoses had been made by professional clinicians. The autism sample comprised 18 men and 5 women, and the control group consisted of 18 men and 6 women, none of whom had received a diagnosis on the autism spectrum at any point. Also, all participants were required to have normal hearing and to be fluent in English (given that word identification was to be a feature of the testing).

Data on the measures of BPVS, AQ, TAS-20, BVAQ-B and age, broken down into groups, are given in Table 3.

Table 2: Mean, (SD), *range* of scores for each group on the TAS, BVAQ, AQ, BPVS measures, and age

Group:	Control (N = 24)	Autism (N = 23)
TAS scores	43.5 (8.9) *30 - 58*	56.6 (10.6) *31 - 72*
BVAQ scores	45.6 (7.0) *29 - 59*	53.0 (11.0) *39 - 77*
AQ scores	14.1 (7.6) *1 - 27*	33.7 (7.0) *23 - 47*
BPVS scores	150.0 (11.7) *120 - 163*	154.7 (9.6) *123 - 165*
Ages	32.5 (13.7) *21 - 63*	36.9 (12.3) *18 - 61*

There was no significant difference in age between the two groups: $t(45) = 1.16$, n.s., and the gender difference was not significant in Fisher’s exact test of ratio inequality, with $p = .533$ (one-sided), n.s. The design of the experiment had been approved by the Goldsmiths Department of Psychology’s Ethics Committee, and all participants gave written informed consent.

This was a convenience sample. A number of the autism group participants had taken part previously in non-musical psychology research conducted by colleagues on the Goldsmiths staff (seven of them had been participants in the study reported in chapter 3 of this thesis). A number of others were recruited through a notice placed with the National Autistic Society. The control participants were recruited through a notice placed in Goldsmiths, and among students and other contacts, both inside and outside Goldsmiths. In all cases it was made clear when seeking participants that no prior musical knowledge or expertise was required as a pre-qualification for taking part. Given that the number of words chosen by participants (an experimental measure, relating to Step 3 in Figure 1) might be expected to depend partly on their receptive vocabulary, and that this might create a false group difference or interfere with the results in other unpredictable ways, the receptive vocabularies of participants was measured using the British Picture Vocabulary Scale (BPVS: Dunn et al., 1997). Whilst the test was developed primarily for testing receptive vocabulary in children, given the level of difficulty in the hardest sections of the test, it appeared sufficient at least to detect any major differences in receptive vocabulary between the autism and control groups. For this reason a truncated version of this test was used, employing the top end of the difficulty range, noted as being suitable for the 16-21 age group. Although there was a slightly higher mean BPVS verbal vocabulary score in the autism group compared with controls (154.7 vs. 150), the difference was not significant ($t(45) = .133$, n.s.).

5.6.3 Cut-off categories for AQ data

A conventional cut-off value for determining likely membership of the autism spectrum is to take a score of 26 or more on the AQ as indicative for screening purposes (eg Sizoo et al., 2009; Woodbury-Smith et al., 2005). When this criterion was taken, and the classification compared with the actual diagnostic category of the participant, the matching was found to be as in table 4 below.

Table 3: Showing number of participants in categories of membership of ASD group and AQ score.

	Control Group	Autism Group
AQ < 26	21 (87.5%)	2 (8.7%)
AQ ≥ 26	3 (12.5%)	21 (91.3%)
Total	24	23

This gives an average accuracy of classification of ASD using the AQ of 89.4%.

5.6.4 Alexithymia: recommended cut-off points

Deborde et al. (2008) list the following points for determining the presence and absence of alexithymia for the TAS-20 and BVAQ-B tests:

TAS-20: total score < 44: alexithymia absent;
total score > 56: alexithymia present
BVAQ-B: total score < 44: alexithymia absent;
total score > 53: alexithymia present.

Individuals with scores intermediate between these two values for a given instrument are considered to have borderline scores, and are not categorized as either alexithymic or non-alexithymic in order to avoid misdiagnosis.

5.6.5 Cut-off categories for TAS and BVAQ data

Using the cut-off categories from Deborde et al. (2008) cited above, the following Table 5 shows the distribution of individuals into these subgroups, broken down into the autism and control groups.

Table 4: Alexithymia/ASD group distribution for TAS-20 and BVAQ-B tests.

	Control (N = 24)	Autism (N = 23)
TAS non-alexithymic	12 (50%)	2 (8.7%)
TAS intermediate	11 (45.8%)	7 (30.4%)

TAS alexithymic	1 (4.7%)	14 (60.9%)
BVAQ non-alexithymic	9 (37.5%)	6 (26.1%)
BVAQ intermediate	11 (45.8%)	5 (21.7%)
BVAQ alexithymic	4 (16.7%)	12 (52.2%)

It was the case for both tests, considered separately, that some individuals in the autism group fell into the "non-alexithymic" category, and some from the control group fell into the "alexithymic" category.

On examining those individuals who fell into the alexithymic categories for *both* tests, none were in the control group. A similar examination of those who were in the non-alexithymic group for both tests showed only one belonging to the autism group (participant 4: see Discussion section for further comment).

5.6.6 Musical experience questionnaire

In order to obtain details of previous musical experience, a musical experience questionnaire was used (reproduced as appendix 2.1, below). The purpose of this information was twofold: to ensure that there were no significant differences between autism and control groups in musical background that might have confounded the group differences and invalidated the results and to provide the possibility of using a rough measure of musical experience as a covariate to help with noise reduction at the statistical analysis stage.

5.6.7 General procedure

Participants were in most cases asked to complete four questionnaires in advance of the testing, namely the TAS-20, BVAQ-B, AQ and the musical experience questionnaires referred to above. In cases where this was not possible, the questionnaires were completed on arrival. Participants were also sent a briefing sheet in advance of the experiment.

Testing took place in a quiet environment within Goldsmiths, University of London. Sound stimuli were delivered through Sennheiser HD 201 high fidelity headphones, from Windows Media Audio files stored on the computer.

On arrival, I confirmed that participants had read the briefing sheet (allowing time for them to do so if this had not occurred), and asked them to sign a consent form.

I then asked (in the case of autism group participants) for details of the nature of their diagnosis, when, and by whom it had been given.

The next step was to establish a comfortable sound level for the individual participant. The participant listened to one of the music tracks on the headphones, with volume at a low level, and the volume was adjusted by the experimenter until the participant indicated that he/she was comfortable with it.

Participants were then informed in more detail about the GSR measurement process, and were connected up to the apparatus. This used a pair of BioLogic disposable snap electrodes P/N 101603, Supplied by Biosense Medical Ltd., connected to the inner surfaces of the first and third fingers of the left hand, connected via a voltage inducer and current measuring apparatus powered by a 9 Volt battery, to a computer program recording resistivity measurements in kilohms at a rate of 2 Hz throughout the experiment; both equipment and program were designed by Rob Davis, a member of the Goldsmiths Psychology Department staff.

Before playing the stimuli and recording the GSR readings, participants were told that during the procedure, which would involve a series of exposures to 30 seconds of sound stimulus alternated with 30 seconds of silence, they should attempt to focus on the music (or sound) while that was being presented, and to make their mind a blank during the periods of silence. Participants were asked to move physically as little as possible during this phase, since muscular activity was capable of producing artefactual GSR responses, and it was suggested that they make themselves as physically comfortable as possible before the experiment began. A small table was provided on which they could rest the hand and arm connected to the apparatus. I suggested (but did not insist) that they shut their eyes during the duration of the GSR experiment. Most participants complied with this suggestion.

The order of presentation of music versus environmental sounds was randomized between participants, so that half of each group received the music first, and half were exposed first to the environmental sounds. However, musical items were always presented as a block together, as were the environmental sounds. Within each music and sound block, the order of presentation of the individual music and sound items was automatically pseudo-randomized, a facility which had been built into the program designed by Davis.

During the GSR experiment, the trace of the conductivity measurements was visible to the experimenter on the computer screen, but not to the participant, so as to avoid any spurious bio-feedback effects. In most cases, as the participant became more relaxed and the ANS activity lowered, the trace showed an increase of GSR resistance to a fairly stable level, which was usually achieved after between one minute and 90 seconds of relaxation and silence. I then initiated the program with a keyboard stroke, which started the first sound stimulus item. Subsequent items were initiated by further manual keyboard presses, timed by myself, after 30 seconds of silence following the conclusion of the preceding sound item.

In view of the physical restrictions involved in the GSR measurements, participants were encouraged to take a short break after this phase of the experiment, and to move around physically to restore circulation and alertness levels (during the GSR phase, some participants became so relaxed, according to their own accounts, as almost to fall asleep).

The entire experiment, after the formalities and preliminaries, always began with the two blocks of GSR measurements. This phase was followed, after a break, by two tasks, the first involving verbal responses to six items of music, and the second requiring responses to a further set of six musical items and the “bundles” of words mentioned in chapter 4.

For the post-GSR phase, I read out the instructions given in the first part of appendix 4.D, and gave the participant the word list in appendix 4.B. The tunes chosen for this part of the experiment, which were numbers 2, 3, 4, 6, 7 and 12 from the list in chapter 4 appendix A, were then played in a randomized order to the participant, through the Sennheiser earphones. Participants responded by ticking the boxes in the appropriate column (column 1 for the first item, etc) during the course of listening to the music and, if they wished, after it was over, to describe the evoked musical emotions. The items were selected manually by myself, based on a randomized order derived earlier using a random number seed provided by Windows Excel. This procedure permitted the timing of the presentations of the successive items of music to be completely flexible, which was necessary given that participants were allowed as much time as they wished to tick the boxes for the word lists for the successive tunes. When participants indicated that they had completed the task for each tune, the next one was played.

After the word lists had been completed, the form was taken from the participant, who was given the list of “bundles”, and the second half of the instructions in appendix 4.4 was read out to them. Participants were asked to respond orally at the conclusion of hearing each item, by selecting that one of the six bundles of words which they considered most likely to be the one corresponding to the descriptors chosen by the pilot group. By way of an additional gloss, participants were told that this exercise resembled the popular TV show “Family Fortunes”, where the task is to guess the most popular answers to a series of questions put to the general public.

Following this part, the BPVS was administered, a debrief form was supplied to the participant and appropriate payment for attendance and/or travelling expenses was made.

5.7 RESULTS

Note that issues concerning the significance of the data presented in this section, and the implications for the hypotheses under consideration, will be addressed in the subsequent “discussion” section.

5.7.1 Descriptive statistics

Basic descriptive statistics comparing control and autism groups on BPVS, TAS, BVAQ, AQ measures and age are given in table 5.1 above. As table 5.1 shows, group means of the AQ, TAS-20 and BVAQ-B scores showed a higher value for the autism group, with the following statistical test results for independent group t-tests:

AQ: $t(45) = 9.2$, $p < .001$.
TAS: $t(45) = 4.6$, $p < .001$
BVAQ: $t(45) = 2.7$, $p = .01$

For the purposes of the study, it was important to investigate responses on the two alexithymia scales. This involved performing correlations between the factors of the two scales. Before examining the correlations that were found in the TAS and BVAQ factors, table 6 below summarises in more accessible form the information contained above about the factor structure and the reported links between them, and a note on which type of alexithymia each factor is intended to measure. The BVAQ factors which correspond to TAS factors are presented alongside those factors.

Table 5: TAS-20 and BVAQ-B factors and descriptions

Type I	TASF1	Difficulty in identifying feelings	BVAQF3	Poor insight
Type I	TASF2	Difficulty in describing feelings	BVAQF1	Poor verbalising
Type I	TASF3	Externally oriented thinking	BVAQF5	Concrete thinking
Type II			BVAQF2	Poor fantasies
Type II			BVAQF4	Poor emotional excitability

Calculation of the bivariate correlations between the separate TAS and BVAQ-B factors yield not only values significant at the .05 level which were expected from previously cited sources and reproduced in the table above, but additional ones. In particular, there are correlations between TAS F1 and BVAQ F1 and F3, between TAS F2 and BVAQ

F1, F2, F3 and F5, and between TAS F3 and BVAQ F1, F2, F4 and F5. Taking only the largest such correlations, it is the case that TASF1 correlates most highly with BVAQF3 ($r = .874$, $p < .001$), that TASF2 correlates highest with BVAQF1 ($r = .770$, $p < .001$), and TASF3 with BVAQF5 ($r = .523$, $p < .001$).

There were no significant differences between the groups in the musical experience measure, with means of 1.9 in the control group and 2.0 in the autism group: $t(44) = 0.37$, n.s. (details on the scoring of the measure are given below).

5.7.2 Replication of results on group differences in alexithymia dimensions

In order to check whether these data replicated earlier results on cognitive and affective alexithymia (Berthoz & Hill, 2005), the cognitive and affective dimensions were computed separately using the BVAQ-B factorial scores: namely, cognitive dimension = the sum of BVAQ factors 1, 3 and 5, and affective dimension = the sum of BVAQ factors 2 and 4. The ASD and control groups differed significantly on the cognitive dimension, in the direction of higher average cognitive score for the ASD group. The means (SDs) were 27.0 (5.5) for the controls, and 31.9 (7.4) for the ASD group: $t(45) = 2.55$, $p = .014$. The groups did not significantly differ on the affective dimension, though there was a strong trend towards a higher score for the ASD group. Means (SDs) were 18.4 (3.7) for the control group and 21.1 (5.6) for the ASD group: $t(45) = 1.99$, $p = .053$.

These results replicated the significant effect of type II alexithymia found in Berthoz and Hill (2005), but the near significant result for affective dimension suggested that type I alexithymia might also be present in the ASD group. However, it appeared likely that this might be an artefact, resulting from a possible overlap in what the affective and cognitive dimensions of the BVAQ actually measured: in other words, from a possible cognitive element in the supposedly purely “affective” dimension. This suspicion was confirmed by the fact that the measures had a substantial positive correlation ($r = .358$, $p = .014$). One method of checking whether the control and autism groups differed on the affective dimension once the common variance shared with cognitive dimension had been excluded, is to conduct a hierarchical logistic regression of autism group membership as the DV against the cognitive and affective dimensions as IVs, with the cognitive dimension entered at step 1 and the affective dimension at step 2. This showed a significant effect of cognitive dimension (p = .022) but the remaining influence of affective dimension did not even approach significance (p = .212).

This view of the essential irrelevance of the affective dimension of alexithymia to this study was reinforced further by the consideration of bivariate correlations, comparing those of the cognitive and affective dimensions of alexithymia with other important variables which would be expected to have a connection with alexithymia, namely wordcount, the two GSR variables, and the Ratio variable. Calculation of the correlations shows that the affective dimension has no significant correlation with these; more specifically, the correlation with Ratio is very small (r = -.076, n.s.), suggesting that sensitivity to the physiologically arousing effects of music appears to depend not at all on the affective dimension of alexithymia. The absence of any correlation between the affective dimension and wordcount (r = -0.009, n.s.) suggests, finally, that the effects of this dimension are negligible on the outcomes to this study. In what follows, alexithymia is considered solely in its cognitive aspects.

5.7.3 Inferential statistics/experimental outcomes

5.7.3.1 Comparison of “wordcount” variable with AQ, TAS, BVAQ.

Bivariate Pearson’s correlations were examined between “wordcount” and other variables of interest, namely TAS-20, BVAQ-B and AQ scores. None was significant, except that the correlation between wordcount and BVAQ showed a strong trend ($r = -0.274$, $p = .062$), but it did not attain significance. When wordcount was correlated with individual factors within the TAS and BVAQ measures, all the correlations were non-significant, with the exception of BVAQ factor 3 (“poor insight”), which correlated at $r = -0.39$, $p = .007$.

5.7.3.2 Group differences in wordcount

The number of words indicated in response to the standard musical items was analysed in several different ways. The first option chosen was to check whether the two groups, autism and control, differed significantly as to the total number of words checked, irrespective of which words they were or to which item of music they referred. The results showed a difference in favour of the control group, with a mean of 25.7 (SD =

7.1) words checked compared with 17.6 (9.1) for the autism group. An independent samples t-test showed that this difference was significant, with $t(45) = 2.46$, $p = .017$.

5.7.3.3 In-depth analysis of wordcount results

I explored this result in more detail, to determine whether it might be due to a difference in the profiles of verbal responses in the case of certain categories of words, or certain items of music. A mixed between/within participants, repeated measures analysis was carried out. The group variable (control, autism groups) was the between-subjects variable, and total numbers of words ticked by individuals for each of the six songs, was the six-level repeated measures variable. This yielded non-significant results for both the main effect of the repeated measures variable of musical extracts (multivariate test: $p = .084$, n.s.) and for the interaction between the music extracts and ASD group variables ($p = .541$, n.s.), though as before in the one-way ANOVA, the main effect of ASD group was significant ($F(1, 45) = 6.34$, $p = .015$).

A similar result was observed when the responses were broken down into the individual words checked, of which there were 28 in all, with the different levels of this variable representing the number of times that word was checked across songs for a particular individual (it could therefore take integer values between 0 and 6 inclusive). There was a significant effect of word on response (ie considering different words as different levels of a repeated measure "words" variable, there was a main effect of "word", ie the average number of times that each word was chosen, taken over the whole sample, differed from one word to another). However, there was no significant interaction between the word and ASD group variables ($p = .891$, n.s.), that is to say, the profile of the average number of times each word was used, when considered for the control group, was parallel to the profile when considered for the autism group.

The profile of the word by group response is shown in figure 9 below. The first 14 words correspond to the list of "words used by the autism group" listed, in that order, in Appendix 4.2. The final 14 words correspond to the "words used by the pilot control group" also listed in that appendix. The lack of word by group interaction is confirmed by the absence of any visible difference between the height of the control line above the ASD group line, when the first and second groups of words are compared.

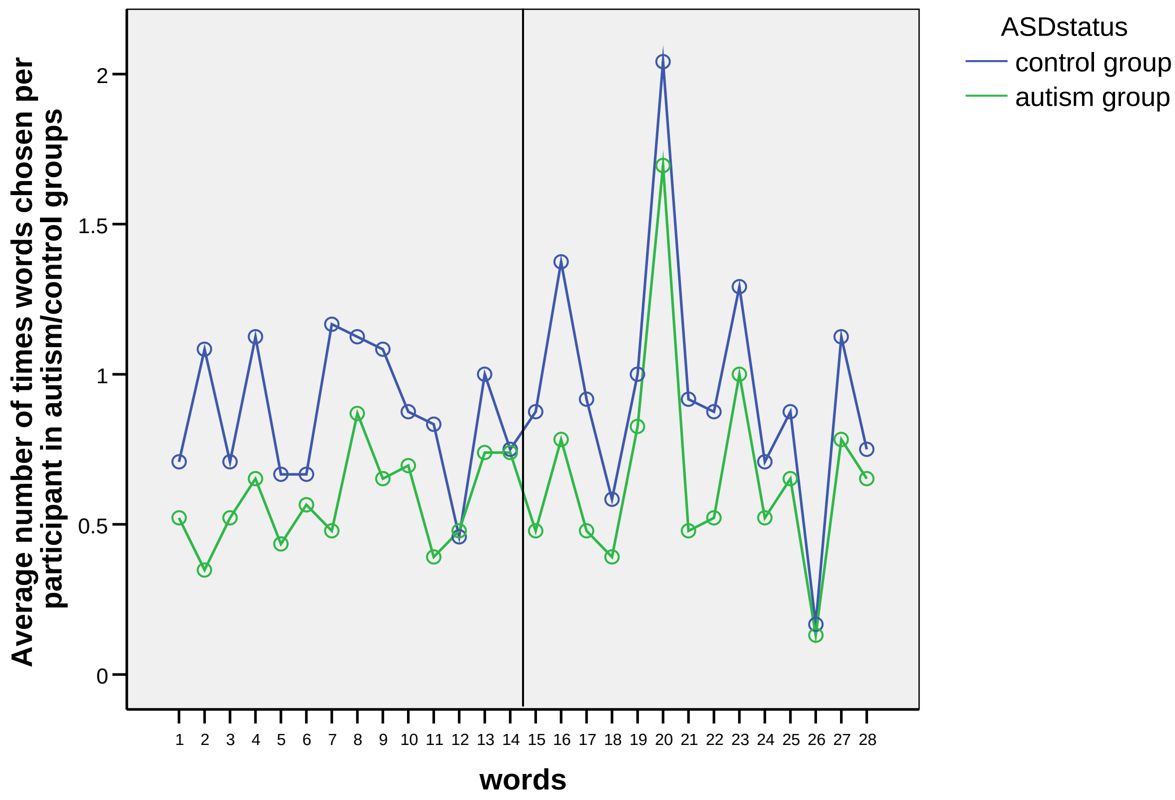


Figure 9: Mean number of times each word was chosen by individuals, broken down into mean values for members of the control and autism groups

5.7.3.4 Internal consistency values

It is possible to treat the individual musical extracts, or the individual words, as items in a test battery, and measuring their internal consistency using Cronbach's alpha. Internal consistency is an important index of the extent to which a set of test items are consistent in terms of the results of what they measure, ie it shows to what extent the items are all measuring the same thing. The result for the musical items is alpha = .887, and for the words, it is alpha = .878, both of which values indicate a high value of internal consistency.

5.7.3.5 Testing for group difference in types of emotion words chosen

One initial hypothesis was that the autism group would tend to choose words reflecting internal arousal states rather than the more externally focused words which were derived from the piloting process described in chapter 4. It will be recalled that the list of words was derived by taking the 14 words used by the autism sample in the chapter 3 study to describe their reactions to music, and adding the most popular 14 words used by our pilot control group to describe their responses to the standard musical items. This gave two categories of words, and it had been expected that there would be a trend in the autism group in this current experiment to prefer the "autism" words over the "control" words, and vice versa.

This trend was tested for by adding the number of words in the "autism" category chosen overall by each participant and subtracting the number chosen in the "control" category. When this number was compared across the groups statistically, the means were practically identical (-1.25 compared with -1.36), and there was certainly no significant difference: $t(45) = .066$, n.s.

Next, the total "wordcount" variable was analysed further to detect the effect of other variables, by means of a multiple regression with Wordcount as the DV and group membership (autism/control) and each BVAQ factor score as the IVs. Since group membership already accounted for a significant effect on wordcount, I entered this first in a hierarchical regression, and the other BVAQ factor scores at the second stage of the regression. The only factor showing a trend towards significance at this step was BVAQ factor 3 (poor insight). See table 9 below.

Table 6: Hierarchical regression of wordcount on ASD group and all BVAQ factor scores.

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	25.750	2.311		11.140	.000
	ASDstatus	-8.141	3.304	-.345	-2.464	.018
2	(Constant)	34.081	8.682		3.925	.000
	ASDstatus	-6.203	3.583	-.263	-1.731	.091
	BVAQF1	.010	.758	.003	.014	.989
	BVAQF2	.576	.602	.153	.958	.344
	BVAQF3	-1.058	.626	-.299	-1.690	.099
	BVAQF4	.143	.666	.034	.215	.831
	BVAQF5	-.610	.732	-.140	-.834	.409

a Dependent Variable: Wordcount

The regression was repeated therefore with factor 3 as the sole variable at the second stage of the hierarchical regression, and it added significantly to the model (p = .024): see table 10.

This table also shows in the results for "model 2", that when BVAQ factor 3 and group membership are both included simultaneously in a forced entry regression for the wordcount DV, group membership shows only a strong trend to significance (p = .065).

The contributions of the variables to the variance in the DV, as measured by the squared semi-partial correlations, also showed, when BVAQF3 is compared with group membership, a greater contribution to variance in the DV (0.097 vs. 0.064), confirming that the effect of group difference on wordcount (i.e. the effect when group is used as the only IV, predicting wordcount) was substantially mediated by that of BVAQF3 ("poor insight"). See values for "part" correlations in table 10 below: in SPSS terminology, part and semi-partial correlations are identical. The squares of -0.252 and -0.312 are .064 and .097 respectively.

Table 7: Hierarchical regression of wordcount on ASD group and BVAQF3, with partial correlation statistics.

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	25.750	2.311		11.140	.000
	ASDstatus	-8.141	3.304	-.345	-2.464	.018
2	(Constant)	36.877	5.252		7.022	.000
	ASDstatus	-6.169	3.263	-.261	-1.890	.065
	BVAQF3	-1.141	.489	-.323	-2.334	.024

Model		Correlations		
		Zero-order	Partial	Part
1	(Constant)			
	ASDstatus	-.345	-.274	-.252
	BVAQF3	-.390	-.332	-.312

5.7.3.6 Analysis of physiological data and construction of “Ratio” variable

The GSR reactions of the participants to the standard set of musical items were analysed. It should be noted that one participant from the autism group proved to have such high levels of skin resistance that no readings were possible. In addition, no environmental noise readings were taken from the first four participants from the autism group. This was because the benefits of having a control “noise” condition only occurred after the testing had begun.

The raw physiological reactions to the music as well as to the environmental noise stimuli, analysed as explained in appendix 5.1, are given in table 11 below.

Table 8: Group mean, (SD), *range* for GSR reactions to music and to noise

	Control (N = 24)	Autism (N = 22)
GSR music	0.65 (0.54) *0.05 – 2.07*	0.29 (0.25) *0.03 – 0.99*
	Control (N = 24)	Autism (N = 18)
GSR noise	0.42 (0.40) *0.02 – 1.79*	0.26 (0.28) *0.03 – 1.05*

While the control group showed a stronger physiological reaction than the autism group to the environmental noise stimuli, this difference was not statistically significant: $t(40) = 1.472$ (n.s.). Moving to the musical stimuli, the reactions in the control group were stronger than in the autism group with $t(44) = 3.00$, $p = .006$.

Significant correlations were observed between the GSR music and noise variables, both for the sample as a whole and within the experimental groups. In the sample, the bivariate correlation was r = .806, p < .001.

In the control group, r = .873 (p < .001), and in the autism group, r = .552, p = .018.

As a measure of the relative reactivity to music I took the ratios between the scores for the music reactions to the 12 items and the noise reactions to the environmental sounds. The new "Ratio" variable showed a slightly higher mean value in the autism compared with the control group (table 14), though this was not significant: $t(40) = 0.57$ (n.s.).

Table 9: Mean, (SD), *range* of scores for each group on the Ratio variable

Group:	Control (N = 24)	Autism (N = 18)
Ratio	2.07 (1.64) *0.54 – 6.98*	2.52 (2.98) *0.22 – 9.79*

5.7.3.7 Regression analyses involving "Ratio" variable

When Ratio was entered as the IV in a simple linear regression of wordcount on Ratio, the regression coefficient for Ratio was significant. When BVAQ-B factor 3 (poor insight) and Ratio were entered as IVs in a forced entry regression with wordcount as the DV, both IVs were significant.

Furthermore, BVAQ-BF3 (BVAQ-B factor 3) and Ratio between them accounted for 25.2% of the variance in wordcount (measured by adjusted R-square). These results are summarised in table 15 below, where Reg 1 refers to the first regression, of wordcount on Ratio, and Reg 2 to the second, of wordcount on Ratio and BVAQ-BF3.

Table 10: Regression of wordcount on Ratio, and on Ratio with BVAQ-BF3

Reg 1		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
	(Constant)	18.712	2.558		7.314	.000
	Ratio	1.631	.800	.307	2.040	.048
Reg 2		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
	(Constant)	34.918	5.474		6.379	.000
	Ratio	1.781	.719	.335	2.475	.018
	BVAQ-BF3	-1.537	.471	-.442	-3.262	.002

Reg 2	R	R Square	Adjusted R Square	Std. Error of the Estimate
	.537(a)	.288	.252	10.514

The SPSS multiple regression program enables the automatic calculation of a "predicted" value of the DV, in this case, wordcount. The predicted value is simply the value calculated using the regression coefficients computed by the program, using the values given for the IVs for each participant. The visual comparison predicted against actual values of the DV is known to be a useful check against outliers caused by faulty inputting of data or other causes. To check for visual evidence of the existence of outliers, a scatterplot was created (Figure 10), labelled by the participant case numbers, of actual against predicted wordcount. Note that the autism group participants can be identified by having case numbers running between 1 and 23, whereas control group participants have numbers extending from 101 upwards.

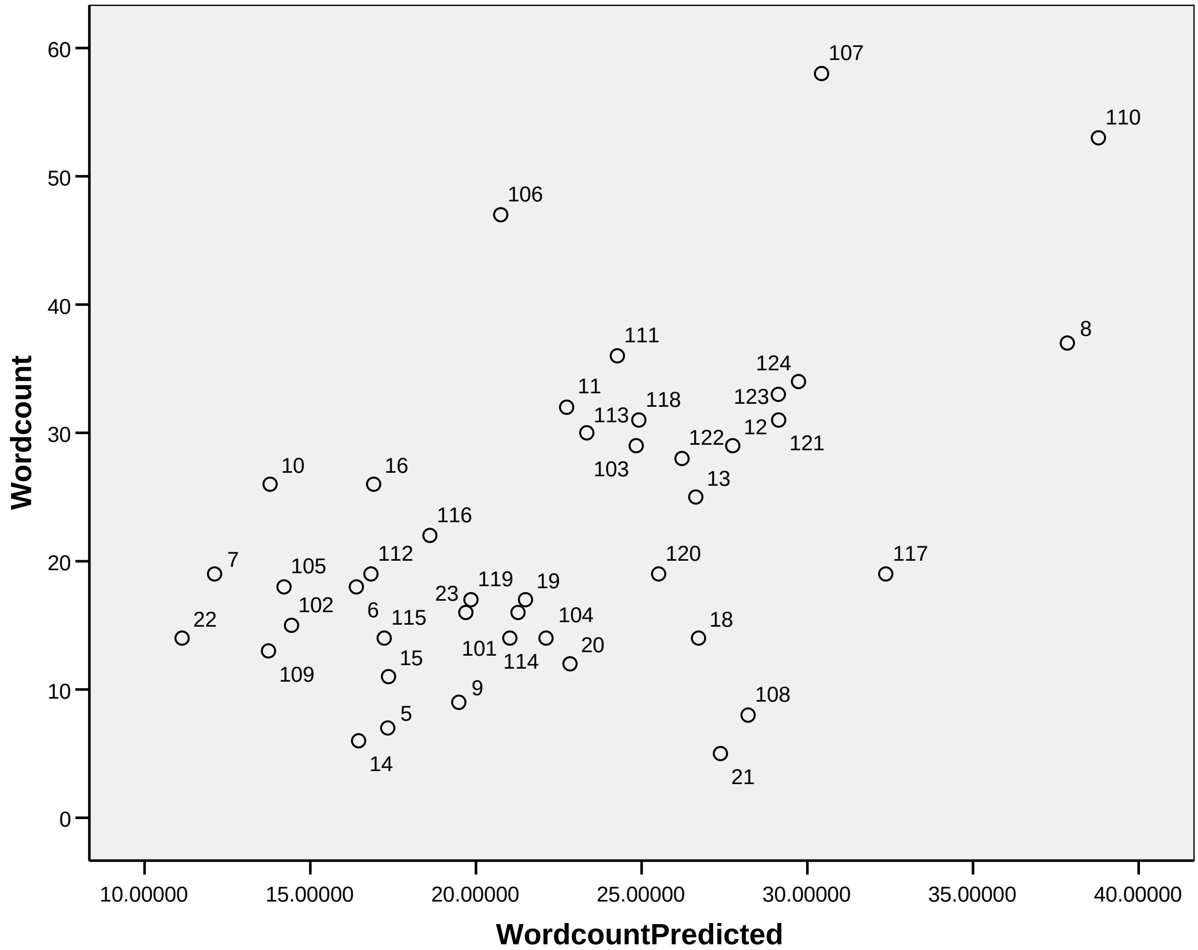


Figure 10: Scatterplot of actual vs predicted wordcount, labelled by participant number

Examining Figure 10, there is no immediate evidence for the existence of outliers. Boxplots for the individual variables confirm this. Case number 107 does lie just outside the boxplot for the wordcount variable, but this participant was also high in physiological arousal (third highest out of the whole group), and testing the Mahalanobis distance of this datapoint from the centroid (8.74) showed it to be well below the level (13.82) which would be significant under the conventional .001 significance level for $\chi^2(2)$. Nevertheless, the fact that case 107 does lie somewhat visibly above the general line of the scatterplot, and cases such as 21 and 108 lie below it, would lead to a desire to find some other causal variable which might help to explain this discrepancy.

In the belief that further information might be discovered by allowing for individual variations in probable liking for the (mainly non-pop) musical items, I coded the musical experience questionnaires on a simple score from 1 to 3 on the basis of years of musical teaching, ability to play an instrument or sing, and expressed interest in classical music.

A forced entry multiple regression of wordcount on BVAQ-B factor 3 and this new "experience" variable yielded significant results for each variable, and an adjusted R-square of 0.213, which therefore accounted for over 21% of the variance in the DV.

I considered that there might be, in addition, a moderating effect of the experience variable and the Ratio variable on wordcount. Moderation is detected by taking the original variables together with their product. It is generally recommended that the variables should be centralised before they are multiplied together, in order to avoid multicollinearity which may be introduced into the IVs if uncentralised variables are used (Dunlap & Kemery, 1987). I therefore created a new variable from the standardised versions of the Ratio and experience variables, namely the product of the Z-scores on the Ratio and experience variables.

When experience, Ratio and the product variable were entered together to predict wordcount, both experience and the product term were significant, although Ratio was not significant. I re-ran the regression with just experience and the product term, and these variables, with an adjusted R-squared of 0.395, now accounted for nearly 40% of the variance in wordcount. With the addition of factor 3 of the BVAQ, these three IVs now accounted for 51% of the variance in wordcount, as measured by adjusted R-squared value. These results are summarised in table 16 below.

Table 11: Regression results for Reg 1 (wordcount on experience, Ratio and product), Reg 2 (wordcount on experience and product) and Reg 3 (wordcount on experience, product and BVAQF3)

Reg 1		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
	(Constant)	13.243	4.260		3.109	.004
	Ratio	-.060	.752	-.011	-.080	.937
	Experience	4.238	1.986	.266	2.134	.039
	Product	6.637	1.638	.571	4.053	.000

Reg 2	R	R Square	Adjusted R Square	Std. Error of the Estimate
	.652(a)	.425	.395	9.454

Reg 2		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
	(Constant)	13.153	4.053		3.245	.002
	Experience	4.220	1.947	.265	2.167	.036
	Product	6.575	1.422	.565	4.623	.000

Reg 3	R	R Square	Adjusted R Square	Std. Error of the Estimate
	.739(a)	.545	.510	8.513

Reg 3		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
	(Constant)	25.559	5.343		4.783	.000
	Experience	4.728	1.760	.297	2.686	.011
	Product	5.822	1.302	.501	4.470	.000
	BVAQF3	-1.234	.388	-.354	-3.179	.003

The scatterplot, with case labels, for the actual and new predicted value of the wordcount appeared as follows, Figure 11.

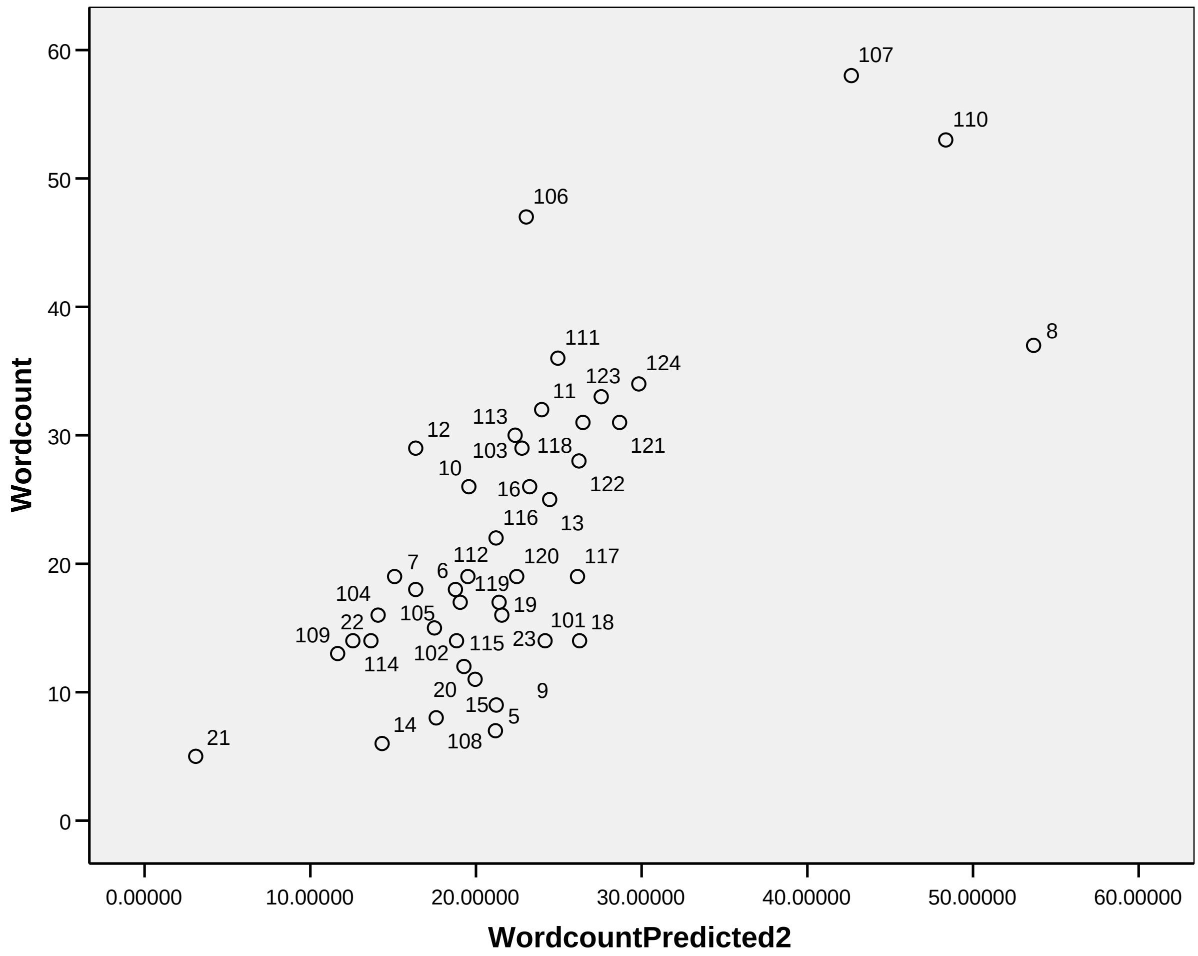


Figure 11: Scatterplot of wordcount against predicted values, labelled by participant number

The appearance of significance for the product term and one of the IVs, but not the other, suggested that the effect of Ratio on wordcount varied with variance along the "experience" axis, from a positive to a negative slope. Such a change would account for both the significance of the product term (as the slope change is diagnostic for a moderation effect), and for the non-significance of the ratio term, the positive and negative slopes cancelling out when taken over all datapoints. This interpretation was checked by carrying out a multiple regression of wordcount on BVAQ3 and Ratio, while simultaneously fixing the values of "experience" at 1, 2, and 3 respectively, where a value of 1 represents no or minimal musical education and experience, 2 represents a more considerable degree of experience, and 3 represents the ability to play an instrument or other evidence of considerable familiarity with music.

This showed the following regression coefficients for “ratio”:

For experience = 1, B = -24.9, *p* = .053
For experience = 2, B = 10.8, *p* = .046
For experience = 3, B = 18.2, *p* = .025.

When a hierarchical regression was conducted with the nuisance variables of experience and the product term entered in the first step, and the hypothesised causal variables of BVAQ factor 3 and the group variable encoding ASD diagnosis in the second step, the following output was noted. Firstly, the model as a whole accounted for 55.5% of variance in wordcount (adjusted R-square = 0.555). Secondly, the “nuisance” variables of experience and the product term accounted for 39.5% out of this total, with an R-square change when BVAQ factor 3 and ASD group were added, of 17.4%, which was significant (p = .001). The squared part-correlations of these two variables showed that the proportion of the unique variance accounted for by BVAQ3 was 0.062, and that of ASD group was 0.053. The greater part of this portion of the variance was therefore due to the BVAQ3 variable, with some remaining variance due to ASD group membership not attributable to the alexithymia component accounted for by BVAQ3.

These results are summarised in table 17 below.

Table 12: Hierarchical regression results of wordcount on experience & product (step 1) plus BVAQF3 and ASD group (step 2)

Model	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
			R Square Change	F Change	df1	df2	Sig. F Change
1	.395	9.454	.425	14.388	2	39	.000
2	.555	8.107	.174	8.019	2	37	.001

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	13.153	4.053		3.245	.002
	Experience	4.220	1.947	.265	2.167	.036
	Product	6.575	1.422	.565	4.623	.000
2	(Constant)	24.715	5.103		4.843	.000
	Experience	4.823	1.677	.303	2.876	.007
	Product	5.976	1.242	.514	4.811	.000
	BVAQF3	-.938	.393	-.269	-2.386	.022
	ASDstatus	-5.957	2.692	-.245	-2.213	.033

Model		Correlations		
		Zero-order	Partial	Part
1	(Constant)			
	Experience	.331	.328	.263
	Product	.596	.595	.561
2	(Constant)			
	Experience	.331	.427	.300
	Product	.596	.620	.501
	BVAQF3	-.420	-.365	-.248
	ASDstatus	-.322	-.342	-.230

5.7.3.8 Predicting ASD membership using logistic regression

I performed logistic regression on the ASD group membership variable as the dependent variable, and various combinations of the covariates as predictors. This is a useful technique for two reasons. Firstly, it is commonly used (in, for example, Berthoz & Hill, 2005) where one wishes to discover whether two groups differ significantly on a continuous variable once the effect of other variables has been eliminated. Since regression looks at the unique contribution of an IV to variance in the DV, logistic regression can ensure this. And secondly, it can provide a reality check, as well as potentially validating new variables as possible clinical predictors of group membership. A variable, or set of variables, which have a high percentage accuracy in predicting group membership, where membership is already known, may plausibly be held up as really telling us something valid and concrete about the condition in question.

As shown in table 18 below, including wordcount on its own was rather poor at predicting ASD membership (59.6% correct) but with the addition of the physiological responses to music, this rose to 78.3%. By comparison, Baron-Cohen's AQ on its own predicted 89.4 of the membership correctly (this happens to be the same result as was obtained above using a crosstabulation table and the conventional cut-off of 26 on the AQ, showing that the conventional value of AQ cut-off is consistent with the data from the present study). When wordcount was added to AQ as a predictor, the accuracy increased to 93.6%. AQ and the GSR music variable predicted membership with 95.7% accuracy, while AQ, GSR music and wordcount together accounted for 100% of the ASD classification.

Table 13: ASD group predicted using wordcount, AQ and GSR music in various permutations.

			Predicted		
			ASDstatus		Percentage Correct
	ASD group predicted using wordcount		control group	autism group	
	ASDstatus	control group	12	12	50.0
		autism group	7	16	69.6
	Overall Percentage				59.6
			Predicted		
			ASDstatus		Percentage Correct
	ASD group predicted with wordcount and GSR music		control group	autism group	
	ASDstatus	control group	18	6	75.0
		autism group	4	18	81.8
	Overall Percentage				78.3
			Predicted		
			ASDstatus		Percentage Correct
	Observed		control group	autism group	
Step 1	ASDstatus	control group	21	3	87.5
		autism group	2	21	91.3
	Overall Percentage				89.4
			Predicted		
			ASDstatus		Percentage Correct
	Observed		control group	autism group	
Step 1	ASDstatus	control group	22	2	91.7
		autism group	1	22	95.7
	Overall Percentage				93.6
			Predicted		
			ASDstatus		Percentage Correct
	ASD group predicted using AQ score and GSR music		control group	autism group	
	ASDstatus	control group	23	1	95.8
		autism group	1	21	95.5
	Overall Percentage				95.7
			Predicted		
			ASDstatus		Percentage Correct
	ASD group predicted using AQ, wordcount and GSR music		control group	autism group	
	ASDstatus	control group	24	0	100.0
		autism group	0	22	100.0
	Overall Percentage				100.0

5.7.3.9 Testing for group difference in "bundles" experiment

The scores on the "Family Fortunes" part of the experiment were as follows: the means for the autism and control groups were 3.9 and 3.3 respectively, though the slightly superior performance by the autism group was not significant ($t(45) = 1.23$, n.s.). See Table 19.

Table 14: Mean (SD) and range of group scores on the bundle identification task

	Control (N = 24)	Autism (N = 22)
Score correct in "bundles" task	3.33 (1.83) *0 - 6*	3.91 (1.34) *2 - 6*

5.8 DISCUSSION

5.8.1 Descriptive statistics

Patterns observed in AQ, TAS, BVAQ and diagnostic group variables

The differences in the control and autism groups regarding AQ, TAS-20 and BVAQ-B are in the direction one would expect given the association of alexithymia with autism, that is, higher scores would be expected on all three measures in the autism group, and this is indeed what was found.

The fact that one individual in the autism group (participant 4) scored within the non-alexithymic categories for both instruments, TAS-20 and BVAQ-B, showed that alexithymia was not invariably associated with autism. The fact that this individual scored at 34 on the AQ, which is comfortably higher than the usual boundary marking the onset of ASD for this measure, suggests that the alternative explanation that the individual was misdiagnosed, can probably be ruled out. This supports data from other research groups showing that alexithymia is not synonymous with autism (Hill et al., 2004; Silani et al., 2008).

The correlations observed between the factors of the TAS-20 and the BVAQ-B are consistent with those reported in the literature, where TAS F1 corresponds to BVAQ-B F3, TAS F2 with BVAQ-B F1 and TAS F3 with BVAQ-B F5. The data in the present study yield additional correlations, but this is not entirely unexpected given that the individual factors within the instrument are expected to be somewhat correlated with one another, since they all measure one or other related aspect of the alexithymia concept. Moreover, the largest correlations for each TAS-20 factor from these data are TAS F1 with BVAQ-B F3, TAS F2 with BVAQ-B F1 and TAS F3 with BVAQ-B F5, which are just the correlations previously noted. The overall pattern of correlations is, therefore, close enough to that predicted from the literature to give some confidence that the tests were properly administered, and were conscientiously answered by the participants (for further discussion of these issues, see Hill et al. 2004; Berthoz & Hill, 2005).

The good correspondence between the diagnostic status of participants and the classification according to the recommended cut-off score of 26 on the AQ suggests that the results are unlikely to have been biased by the existence of mistaken diagnoses in the autism group or undiagnosed individuals with autism in the control group. The diagnostic categories were therefore accepted as they stood, and no attempt was made to eliminate participants from the experiment on the grounds that they fell outside any conventional thresholds for the AQ, nor, for that matter, on the TAS-20 or BVAQ-B. There is a case to be made for excluding such individuals, but given the small numbers in this sample it seemed advisable to retain all participants, at least provisionally, unless and until it became clear that they might be outliers in terms of one or other of the behavioural or physiological variables that I had measured myself. This decision was reinforced by a certain scepticism as to the wisdom of relying totally on self-report measures such as the AQ, TAS-20 and BVAQ-B, an issue which is discussed at some length in chapter 6.

5.8.2 Inferential statistics

Identification of "wordcount" variable with "poor insight" construct

The pattern of correlations observed between wordcount and AQ, TAS, BVAQ and the individual TAS and BVAQ-B factors suggests that the construct being measured by "wordcount" is associated quite specifically with one aspect of alexithymia, namely that factor measuring "poor insight" or BVAQ-B factor 3. Examination of the four questions in the BVAQ-B that load onto this factor lend credibility to this conclusion. The four items are as follows:

"I do not know all that is going on in my mind".
"When I feel lousy, I know whether I am afraid or dejected or sad" (negatively scored).
"When I am fed-up, it remains unclear to me whether I am sad or afraid or unhappy".
"When I am in a sunny mood, I know whether I am enthusiastic or cheerful or elated" (negatively scored).

It is clear from this, that the items collectively measure the ability to put verbal labels of some degree of complexity on states of altered mood, which is precisely the sort of task demanded of participants in the tick-box part of this experiment. Conversely, the strong correlation between BVAQ factor 3 and the wordcount variable reinforces the conclusion that wordcount does after all have some validity as a measure of general emotional insight. Given the arguably special nature of the emotions induced by music, this in itself is a useful conclusion.

It might be objected that the factor of "poor verbalizing", or BVAQ factor 1, should be an equally good if not better predictor of performance on the wordcount measure. However, in this experimental setup the participants are cued with all the words that they need to have, or can use, to describe their feelings and reactions. Therefore it is not so much the ability to find the words for feelings, as the ability to look into one's own mind and connect the feelings one is experiencing with the words on the list that was important; and from examining the relevant factor 3 items (see above), it is evident that factor 3 has a strong degree of face validity in measuring precisely this ability.

5.8.2.1 Group differences in wordcount

The between-groups analysis of the wordcount variable shows that the groups do indeed differ in the direction that one would expect, given that they differ in the measures of alexithymia. The autism group is higher in both measures of alexithymia (TAS and BVAQ-B) than the control group, and the group also ticked fewer emotion boxes overall than the control group. This provides some confirmation that the "tick box" task is indeed measuring a construct which is the same as that measured by the alexithymia questionnaires. The fact that the autism group was in no way inferior to the control group in receptive vocabulary, rules out the possibility that the reduced mean of wordcount in the autism group was due to a simple lack of understanding of language. It was important to be able to rule out this possibility, given that reduced verbal IQ, or at least, uneven verbal/non-verbal IQ profiles, are commonly found in autism.

On examining individual words, it is seen that in every case, the number of times that word was ticked by the control group lies above that for the autism group, and this uniformity of relative frequency suggests that the tendency to tick fewer boxes in the autism group was not much affected by the nature of the individual words available. This visual impression is confirmed by the fact that the interactions between the word and ASD group variables, and between the music extract and ASD group variables, were non-significant. The Cronbach's alpha calculation further confirms this; a value of Cronbach's alpha over 0.7 is usually taken as satisfactory in terms of internal consistency, and a cut-off of 0.8 for a psychometric instrument is taken as being quite strict. The fact that the values obtained in this study approach 0.9 suggests that the tick-box procedure has a high degree of internal consistency, and that there is little variability within the population of music extracts, or words, that were selected to measure it, in the following sense: the probability of selecting a given emotion descriptor is greater for the control group than for the ASD group, but the amount and direction of the difference in probabilities does not appear to vary from extract to extract, or from word to word. That is, the greater ability, or willingness, to choose emotion words to describe their experiences does not depend on the particular words within this sample of words, and therefore one would reasonably infer that the tendency is likely to be characteristic of emotion words in general, and not simply an artefact of the necessarily limited range of words employed in this particular study. To this extent, the study has external validity as far as word choice is concerned.

This does not of course imply that the musical items that were chosen necessarily have the same degree of generalizability or external validity. The musical items did not in any way attempt to be a random selection from the total population of items that might have been chosen. It is possible, indeed plausible, that items selected from a different population, such as heavy metal or other types of non-classical music, might give rise to different results. However, it does suggest that within the genre of music from which the extracts were drawn, the items showed a high degree of homogeneity.

5.8.2.2 Lack of group difference in types of emotion words chosen

With regard to the types of words chosen by the groups, in terms of “autism-type” and “control-type” words, it had been expected that the groups would exhibit a difference in the relative numbers of these types of words chosen. Specifically, if “autdiff” is defined for each participant as the difference between the total number of words chosen by that participant which belonging to the “autism-type” wordgroup, and the “control-type” wordgroup, “autdiff” had been expected to differ significantly between the two groups. In fact this variable had been thought likely to be higher in the autism group, reflecting a hypothesised greater tendency in the autism group to choose internally focused words to describe their reactions, compared with the control group. However, there was almost no difference at all observed. This part of the hypothesis was not confirmed. It might be argued that one reason for this might be that when the autism group in the previous study were asked to describe their reactions to music, and produced a set of internally-focused words such as “excitement”, “calmness” and so on, they were not cued with any other words, and had to give spontaneous replies. When presented with the complete wordlist in the present study, participants with autism were able to respond in terms of externally focused words because they were provided with the cues in the form of the words themselves. It should be noted that seven out of the total of 23 autism group participants had also been interviewed in the previous study described in chapter 3, so that it might have been expected that there would be a tendency for these individuals to use the internally focused words that they had employed before in the different circumstances of that study. The absence of any effect in the present experiment despite this fact, suggests that the change in presentation in this study made a radical difference in the way they responded.

5.8.2.3 Motivation for using the “Ratio” variable

The outcome of the statistical analyses given above suggests firstly that step 1 (physiological arousal) in the model cited in the introduction differs interestingly between the autism and control groups. There was a significantly lower physiological response in the autism group compared with controls. However, when the response to environmental sounds was included in the analysis, this response also showed a lower level in the autism group compared with controls. True, the difference did not attain significance in the case of environmental sounds, but this is due to the larger variance in the response to the environmental stimuli compared with musical stimuli: this conclusion is supported by the fact that when the ratio between the two physiological responses is taken, the means of this ratio are almost identical in the two groups, so that the reduced response to music in the autism group is almost completely cancelled out by the equivalent effect in their response to environmental sounds.

Both music and environmental sound evidently have a physiological effect on listeners. If one considers the effect of music, one can partition it into (a) a component due to the aesthetic and emotional content of the music, and (b) an effect due purely to the impact which noise of a certain decibel level has on the listener, simply by virtue of the fact that human beings have evolved to attend to changes in the general sound environment for obvious reasons of survivability. Component (b) will be affected in ways which are not of direct interest to this experiment, but which amount to individual variations in sensitivity to sound per se.

Differences in sensitivity, if they exist, will affect the physiological responses to both music and noise, because both involve sound of a certain volume. The fact that correlations were observed in both the autism and control groups suggests that a common factor, which can only be due to reactions to sound *per se*, does indeed exist.

It is reasonable to argue that these differences in sensitivity can be factored out of the measurements by dividing the music response by the noise response, which was the solution that I adopted here. This brief account admittedly glosses over certain possible complications, which are more fully addressed in chapter 6.

5.8.2.4 Regression results involving “Ratio” variable

The “Ratio” variable taken on its own significantly predicts wordcount, with a regression coefficient of 9.8. When BVAQF3 is included together with ratio, the model improves in terms of its predictive ability (as measured by adjusted R-squared, the model improves by a significant 19.4%), but the regression coefficient is still close to the previous value, at 10.7.

The significant result for the regression of wordcount on BVAQF3 (poor insight) and Ratio shows that both the ability to identify one’s emotions verbally, and the actual level of physiological arousal experienced, feed into the output of emotion words provided. This makes sense, in that the degree to which listeners are “turned on” by listening to different items of classical and semi-classical music is bound to depend to a large extent on personal musical taste and experience. The physiological arousal generated by the music is likely to depend monotonically on the extent to which listeners are emotionally moved by the music; and the extent of emotional arousal as experienced by the listeners is likely to affect the number of words that they use to describe that arousal, assuming – as seems plausible – that more intense emotions are likely to yield more detailed information as to their nature and content. Just as it is a common experience to notice more detail about a landscape when it is seen in broad daylight compared to its appearance in moonlight, it is plausible that, within limits, a more intense emotional experience will give rise to more detailed and discriminable emotional descriptors than a less intense one.

The inclusion of the experience variable and the moderating factor based on the product of Ratio and experience, not only improves the power of the model (raising the percentage of variance accounted for from 25.2% to 51%), but also, in the scatterplot, shows the former near-outlier cases of 21, 107 and 108 now lying well inside the central area.

However, the introduction of the experience variable, which attempts to discriminate between individuals who are likely to be responsive to classical music (those with experience = 3) and those unlikely to be responsive, shows that this rather bland result hides something much more interesting. For individuals with moderate or high values of experience, the regression coefficients are both positive. However, at the lowest value of experience, ie at experience = 1, the coefficient is strongly negative, at -24.9.

Why might this be? One plausible reason might be that a higher physiological response in those who like classical music is associated with greater cognitive response. Furthermore, those individuals who are antipathetic to classical music might find the musical items that I used (which are mostly strongly based on classical models) somewhat unpleasant or tedious to listen to, and so the stronger physiological reactions are associated with dislike or aversion, rather than liking. Therefore, since the alternative emotion words from which they can choose do not allow them to use words indicating dislike or boredom with the music itself, their reaction will tend to be a reluctance to indicate emotional reactions at all. This speculation requires further investigation using a modified version of the paradigm developed for the current study.

5.8.2.5 Logistic regression to predict ASD membership

Logistic regression has been used as a means of controlling for possible confounding effects of other variables in between-group analyses (e.g., Berthoz & Hill, 2005). Here, it is more useful as a means of telling whether the variables measured in this experiment have any potential diagnostic capability that might add to the already very effective AQ screening instrument. The fact that the AQ predicts membership as well as it does could have been deduced from the previous result that taking a cut-off point on the AQ of 26, 89.4% of participants are correctly classified according to diagnostic category.

Of perhaps more interest is the fact that the raw physiological responsiveness to music enables us to discriminate with 100% accuracy between the autism and control groups. The effectiveness of the inclusion of the music response reflects the fact that as already noted, the autism participants generally reacted less strongly to the music than the controls. This effect appears to have accounted for the remaining variance in group membership unaccounted for by the AQ measurement.

I leave open the extent to which this effect might be clinically exploitable so as to improve the accuracy of the AQ questionnaire as a diagnostic tool. It would be unsafe to generalise from this relatively small sample to the wider ASD population, especially so given that the real cause of the group difference in reaction to music was not clear from this experiment (but see also chapter 6).

5.8.2.6 Lack of group difference in "bundles" experiment

The lack of significant difference between the groups in the "Family Fortunes" part of the experiment suggests that the purely cognitive-to-verbal component of the emotion naming process, symbolized as step 3 in the model given above (Figure 5.12), is normal in autism in the conditions of the experiment. Incidentally, one could comment that if the ability to imagine how other people are likely to react emotionally to a musical stimulus is any measure of "theory of mind", than the autism group showed no deficiency in theory of mind compared with the control group.

This might appear odd, especially to those who are inclined to adopt a very simplified version of the theory of mind theory of autism, in which it is maintained that it is characteristic of autism to evince a deficit in theory of mind, which may be so extreme as to be appropriately termed "mindblindness". While it is the case that children with ASD are generally impaired in theory of mind tests, improvement in skill occurs with development to the point that many children are able to 'pass' at least simple theory of mind tests within experimental settings, and theory of mind ability has been shown to relate to verbal mental age in those with ASD (e.g., Happé, 1995).

In terms of data from adults, studies specifically designed to discriminate between first and second-order theory of mind abilities in adults with autism have found both intact first-order abilities in the autism group, but also a group difference in tasks involving second order abilities (e.g., White et al., 2009). This is consistent with the interpretation that the "bundles" task as presented here may be simple enough (being essentially a first order task) to present no difficulty to intelligent adults, who may have learned formal strategies for correlating external stimuli such as music with external emotion words. It should, however, be noted that even in the White et al. study, although there was a significant group difference in performance, with the ASD group performing significantly worse than matched peers on a theory of mind task, the data still indicate that some 70% of adults in the autism group responded accurately on a second order theory of mind task. Therefore, *group* differences cannot reliably be generalised to all adults *within* the group, and the evidence is consistent with the supposition that some adults with autism perform well on theory of mind tasks, at least within the constraints and artificialities of an experimental setting.

5.8.3 Implications for the model

Returning to consideration of the model presented in Figure 8 (and elaborated on in Figure 12 below), the findings reported in this chapter suggest that step 1 is of a similar magnitude in the autism group to the control group, assuming – as has been suggested above – that the correct measure of arousal corresponding to this step is the Ratio variable. The similar performance on the “bundles” test suggests that step 3 is also intact in autism. The fact that the outcome of steps 1 to 3 – ie the verbal report of emotional arousal – differs between groups, therefore implies that there must be a difference between the two groups in step 2. This result, and possible alternative explanations for these findings, will be further discussed in the following chapter.

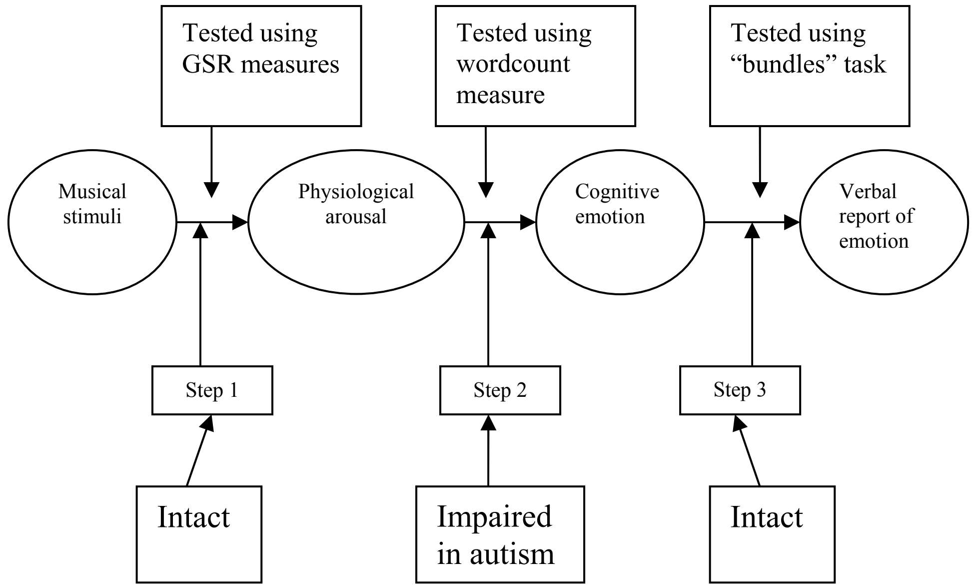


Figure 12: Initial model, and results from testing of it.

Chapter 6. Summarizing the ideas of the thesis, retrospectively and prospectively

Folly consists in the desire to reach conclusions.

Gustave Flaubert

6.1 SUMMARY

In this chapter, the overall conclusions of the thesis will first be stated. Then some methodological issues relevant to the study reported in Chapter 5 will be made, followed by a discussion of how the findings of Chapter 5 feed into the model presented in that chapter. The current chapter ends with proposals for possible clinical applications for the ideas developed in this thesis.

6.2 OVERVIEW OF RESULTS

If it were true, as urged by some philosophers of music, that musically evoked emotions did not exist, then an investigation comparing the effects of music on the emotions in autism and the typically developing population would be doomed to failure from the outset. The experiment described in chapter 2, which is most plausibly interpreted as showing an effect of music in altering underlying mood states, is intended to show that this objection cannot be universally true. The results indicate that besides music's known aesthetic impact and its ability to induce strong feelings of pleasure ("chills") in some individuals (Salimpoor et al., 2009), music can also have a more profound effect on affective state. The influence of musical extracts of different moods on behavioural choices in a face/mood identification task suggests that there is an underlying emotional response in listeners, at least at the basic level, which can be interpreted in terms of a simple happy/sad mood distinction.

The possibility that this is a mere verbal priming effect is made less plausible by the existence of a negative correlation between musical experience and the size of the effect on individuals, which is difficult to explain on the basis of any known priming mechanism, but which is understandable when the pattern of higher left-brain activation typically found in expert musicians is taken into account.

Having established that at least on a basic level, genuine emotions do appear to be evoked by music, chapter 3 examines the assumption that individuals on the autism spectrum fail to respond emotionally in the manner of the typically developing population. By means of a qualitative method involving semi-structured interviews and the analysis of the data using Grounded Theory, it was shown that high-functioning adults with autism or Asperger Syndrome do, plausibly and consistently, report being emotionally moved by music. The testimony of a number of members of this sample that they have used music for “therapeutic” purposes, and to manage their moods on a day-to-day basis, lends credibility to the conclusion that the affective impact of music on individuals on the autism spectrum is a genuine one, and not a spurious effect which is in fact due to confabulation, confusion between metaphor and reality, or a concession to social desirability characteristics of the study. The results also suggest that although the autism group showed many points of similarity with other research published on typically developing samples, it also demonstrated a degree of impoverishment in terms of the range of words used to describe the emotional effects of music.

The results of chapter 3 suggest that in mapping the differences between ASD and TD individuals in terms of their emotional responses to music, it is worth focusing on the differences in the language that they use to describe their experiences. It is hypothesised that the verbal differences may be due either to a lack of emotional reactivity (i.e., type I alexithymia) or to a lack of cognitive access to the nature of their emotions (i.e., type II alexithymia), already reported in the literature. Chapter 4 describes the work undertaken to develop a suitable set of verbal stimuli to be used, in conjunction with a set of matched musical and environmental noise stimuli, to elucidate the extent to which an ASD group uses emotion words less than, or differently to, a TD sample.

Chapter 5 brings together the previous threads of the research by exposing matched TD and ASD groups to a standard set of musical and environmental noise stimuli. Here, the aim has been to demonstrate that when proper allowance is made for abnormal sensitivity to sound per se, the physiological response to music in the ASD group, measured using GSR, is no different from that in the TD group, indicating that there is no evidence for type I (affective) alexithymia in ASD. However, differences noted in the number of words used by the ASD and TD groups to express their evoked emotions can be largely explained by differences in one particular measure of type II alexithymia, that measured by factor 3 ("poor insight") in the BVAQ-B questionnaire. It is argued that this factor is the one in both the TAS-20 and BVAQ-B questionnaires which most plausibly measures the ability to achieve detailed insight into one's internal emotions.

Chapter 5 also confirms that the previously observed differences between the autism and control groups, showing a higher value of type II alexithymia in the autism group but no significant difference in type I alexithymia, persist in the sample chosen for this experiment.

6.3 METHODOLOGICAL ISSUES

6.3.1 The possible value of GSR/wordcount-based measurements of alexithymia

It is proposed in this section that self-report measures introduce validity issues when it comes to evaluating alexithymia in comparisons between control and ASD participants, and that this is perhaps more the case than in comparisons not involving ASD. It is suggested that GSR-based readings of arousal, coupled with a wordcount task which does not require a direct self-evaluation of alexithymic tendencies as do the TAS-20 and BVAQ-B tests, might provide a reliable, objective measure of alexithymia which would avoid the social desirability characteristics of the questionnaire-based evaluations.

6.3.2 The pervasive nature of self-report bias

Questionnaire-based measures in psychology can be divided into those that involve self-report of traits and attitudes, and those which depend on tests of performance. The latter type is arguably impossible to fake, except of course in the direction of doing worse than one's capabilities allow, an outcome by which the participant is seldom motivated. But self-report questionnaires which attempt to measure personality characteristics can undoubtedly be manipulated by sufficiently knowledgeable and Machiavellian participants, to give a misleading impression. This vulnerability of self-report measures to "social desirability bias" and other threats to validity is a well-known phenomenon, and has received some considerable attention from the profession, at least theoretically, but recent research suggests both that it is pervasive, and that insufficient steps are generally taken to guard against it (Adams et al., 1999; Stuart & Grimes, 2009; van der Mortel, 2008).

6.3.3 Self-report bias compounded by failure of insight

In these circumstances, the use of self-report questionnaires such as the AQ, TAS-20 and BVAQ-B to measure what are essentially deficiencies and inadequacies, must consider that a social-desirability bias may be a factor in the answers that are given. But in the case, especially, of the alexithymia measures, another factor comes into play. An alexithymia questionnaire aims at measuring the extent to which an individual has insight into, and understanding of, their own internal emotional states. The validity of the measure depends on the ability of the participant to provide accurate answers to a series of questions about their internal mental states. If that ability is compromised, the answers to the questionnaire may be inaccurate: in other words, the validity of the measurement is guaranteed only for those participants who score near the bottom end of the range for alexithymia (i.e., have the best insight into their emotions).

As one example, take item 11 in the BVAQ: "I can express my feelings verbally". This is supposed to measure factor 1, "poor verbalizing", and its face validity is clear. But a person with no insight into, or knowledge of, emotion words beyond the basics may well respond with a "strongly agree" to this question. A person of great insight and complexity may, on the other hand, respond with "strongly disagree", on the grounds that it would take the verbal dexterity of a Proust to do justice to the complexity of their internal mental life.

As a check on this, I carried out multiple regressions on the total alexithymia scores (for the two scales separately) to establish whether there was any correlation between these scores and other variables, once ASD group membership had been partialled out. The presence of such a correlation might suggest that this variable was acting as a noise variable, or a confound, influencing the self-assessment score in ways which appeared inconsistent with the validity of the alexithymia measures.

6.3.4 How can gender influence TAS score?

The variables tested in this way were age, BPVS score and gender. The only significant result observed was that for the TAS-20, where there was a small, but significant effect (p = .049) of gender, with women on average scoring 6.5 points higher on the TAS-20 than men. There was a weak trend, though not attaining significance, for women to score higher than men on the BVAQ-B ($t(45)$ = 1.5, p = .136), with the mean for women at 4.8 points higher than that for men. This trend could not have been due to a gender imbalance in the groups: there were five women out of 23 in the autism group, and six out of 24 in the control group.

Why might this difference have occurred? At this retrospective stage one can only speculate. It is of course possible that the effect is a genuine one, and that the women in the sample just happened to be more alexithymic on average than the men. But it is at least as likely that those women with a degree of alexithymia were simply more *honest* about their self-perceived lack of emotional discernment than the men. It might also be that although lacking emotional insight, women with alexithymia might be possessed of another form of insight – insight into their disabilities – and in this case, insight into the extent of their alexithymia. It is also possible that the difference, which was only just significant, was simply a statistical artefact. Other studies (e.g., Lane et al., 1998) have found a small tendency for a higher level of alexithymia in men than in women: though see also Haviland, Warren and Riggs (2000), who found an effect in the opposite direction.

An example will be given of an individual whose self-assessed scores on the AQ, TAS-20 and BVAQ-B did not, in the author's opinion, reflect the true state of affairs. A male participant in the autism group, MF, scored 26 on the AQ, a point which is generally taken as the boundary between normal scores and the start of the autism spectrum. In dealing with him on a personal basis, it became clear to the author that he possessed perhaps the most extreme example of autistic traits that the author has encountered (this was not due to low intelligence: his BPVS score, at least, was only half a standard deviation below the average for the group). This person's TAS-20 and BVAQ-B scores were also below the limits at which he would be declared as clearly having alexithymia

– they both lay in the intermediate range, in which it is advised that no clear diagnosis of alexithymia can be made.

I conclude that useful though self-report questionnaires may be as a means of gaining a great deal of data quickly and easily, they share all the faults noted in other instruments of this type. This demonstrates the benefits of finding some other procedure which can take some valid measure of the constructs involved. It may be that alternative measures of alexithymia, involving a combination of the verbal and physiological predictors developed in this thesis, may suggest a way forward for developing a more reliable and objective measure of alexithymia, which may be far more resistant to demand characteristics such as social desirability bias. Further study of this problem would need to evaluate this proposal in conjunction with other, more objective but traditional measures of alexithymia, such as the OAS (Observer Alexithymia Scale: Haviland et al., 2000).

6.3.5 Possible reasons for group differences on music GSR scores

In Chapter 5, the ASD group showed significantly lower GSR music responses in comparison to their typical peers when exposed to identical extracts. As described in Chapter 5, the ratio between responses to music and noise was calculated, and a justification for this was provided on theoretical grounds. Since the aim is to measure emotional responses to *music* as such, and not simply responses to sound stimuli of a certain decibel level applied to the ears, the clear difference in raw response to music between the control and autism groups demands some explanation. This discussion is placed here, rather than in Chapter 5, since although it is pertinent for the development of the model first presented in Chapter 5 (see Figure 8), the ideas it contains are the result of somewhat speculative deductions from the results derived in the course of the experiment, which go beyond the usual limits of discussion sections and which would make the chapter unwieldy.

6.3.6 Varying sound volume: a possible confound?

One feature of the experimental setup was that I adjusted the volume of the sound stimuli to be comfortable for each individual participant. While this might not be considered a valid approach in some domains as it does not ensure reliability and replicability by maintaining the sound level presented to different participants at a level which is identical in decibel terms, the latter is not wholly appropriate when dealing with individuals on the autism spectrum.

6.3.7 Hyperacusis in autism

It is well known that hypersensitivity to sensory stimuli, and in particular to sound (hyperacusis), is common in autism (Gomes et al., 2008; Khalfa et al., 2004). This means that any predetermined standardised sound level would be likely to be either unduly quiet for control participants, or unduly and unpleasantly – perhaps even painfully – loud for some participants within the autism group. Participants with ASD fall for ethical purposes into the category of being members of a "vulnerable" population, and their comfort and wellbeing must be considered with especial care, and given every consideration. It was therefore decided to allow each participant to set the optimum level of sound themselves, by playing one of the louder music tracks and adjusting the volume control on the computer until the participant indicated that they were satisfied with the result.

6.3.8 Is it preferable to arrange for similar subjective, rather than objective, loudness measures?

Given that the prevalence of hyperacusis in autism has been estimated to range between 15% and 100% (Gomes, 2008), it was very likely that the volume levels chosen by the autism group would have been substantially lower than for the control group (I did not, and with hindsight this was an unfortunate omission, record the precise volume settings for each participant). However, it might be claimed that such volume readings would have had little subjective validity in any case. Although there are obvious philosophical difficulties with comparing the quality of the subjective experiences of two different individuals, it is plausible that for a person with hyperacusis, the subjective sensation of listening to music at low volume is similar to that of a normal person listening to the same music at a higher volume. Therefore, it could be argued, that in terms of the clearly desirable aim of presenting a set of stimuli having the same subjective effects, it was more correct scientifically to let participants choose their optimum sound levels themselves, than to impose upon them a predetermined volume level, even had the latter alternative not effectively been ruled out by the need to cater for every possible cause of discomfort or distress when dealing with participants from a vulnerable population.

In view of these considerations, it is arguable that from the point of view of obtaining behavioural measures of the emotional impact of music, the procedure of using self-selected “comfort levels” of sound volume was not only ethically inevitable but also scientifically justifiable. In other words, measuring emotional reactions using the “tick box” method in response to these sound stimuli presented using this procedure, as a measure of emotional responsiveness to music, was arguably quite valid.

6.3.9 Possible independent effects of subjective and objective loudness levels

But now consider the impact on the GSR, or physiological, aspect of measurement. There are two results to be explained here, which are addressed together. The first one is the lower absolute value of the response in the autism group, a difference which reaches significance in the case of the music stimuli, and is apparent as a statistical trend, in the case of the noise stimuli. The second problem is to explain how this group difference is apparently nullified when the ratio between the music and noise stimuli is taken. Another way of posing this question is to ask how it is that both the music and the environmental noise reactions are reduced in autism, a reduction which can be seen to be effectively identical when the ratios are taken and the two reductions are allowed to cancel each other out.

One model which might satisfy the data obtained would be illustrated in the following modification of the model (figure 13) originally proposed in chapter 5 as Figure 8.

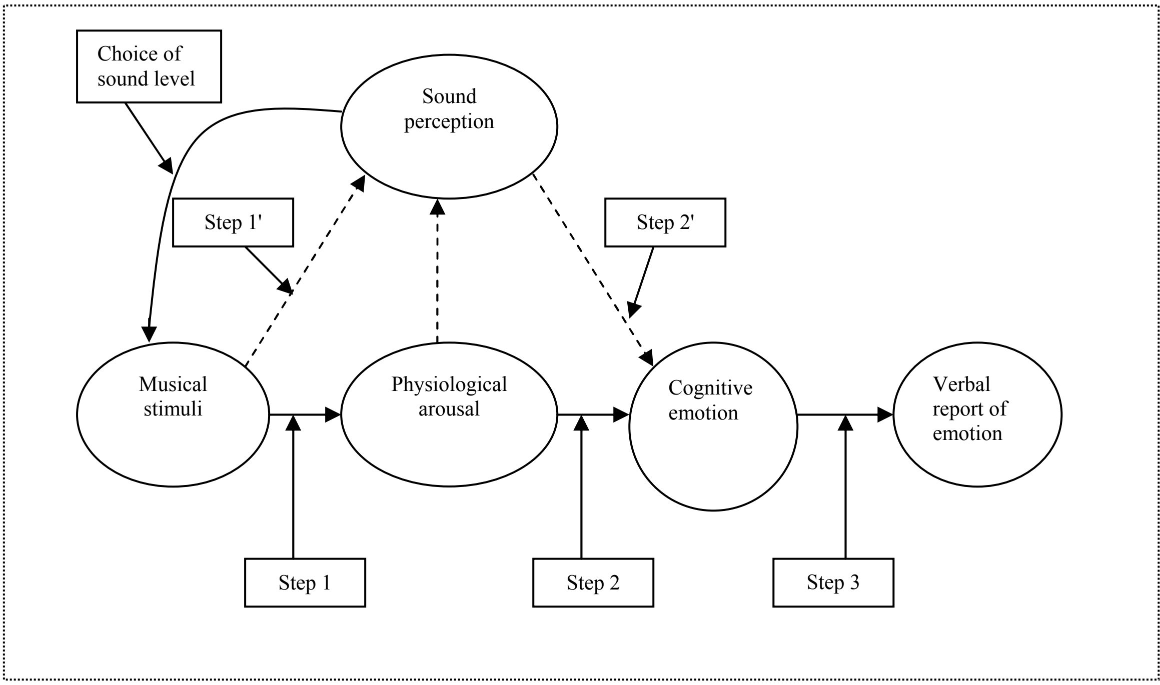


Figure 13: Amended model for effect of music on use of emotion words.

In this model it is hypothesised that there is some stage at the process of perception, labelled simply “sound perception”, at which some judgement of sound volume is possible. Furthermore, this judgement is unusually sensitive in the case of autism. There needs to be a link from the incoming musical stimulus to this measurement process. Assuming this to be direct, the putative link has been labelled ‘step 1'’ in the diagram. However, it is possible also that it might function via the physiological arousal mechanism. Reflecting the very tentative nature of this mechanism, broken rather than solid arrows have been used to indicate it.

The effect of this step is to feed back into the strength of the incoming stimulus, in the case of this experiment, by requesting a reduction in sound volume (an effect which is recalled to have occurred in the case of a substantial minority, perhaps a third to a half, of the autism group, compared with a much smaller number of the control group, though the details were not noted at the time). This in turn will reduce the physiological arousal level mediated through step 1. However, if it is assumed that the sound perception level is now optimal, the actual effect of “sound perception” on upstream elements of the perception and evaluation of the incoming signal, may be subjectively comparable with individuals who do not have a tendency to hyperacusis, and who choose to listen to the music at a higher volume. The lower volume of the music chosen is likely to lead to a lower physiological arousal, assuming that lower volume levels tend to correspond, across different individuals, to reduced physiological arousal. There seems to have been little research investigating this point, though something akin to this phenomenon has been demonstrated in animal experiments (Borg, 1977). However, if this can provisionally be accepted as a reasonable assumption, then the lower volume regarded as optimal by the autism group would lead to a lower level of physiological arousal through step 1, which is consistent with the results actually found.

However, at this point a possible objection to the earlier conclusions might be put forward. Under the original model, the lower sound intensity which, it has been admitted, might possibly associated with the autism group, might account for the reduced wordcount recorded from this group. For if the input to step 1 in this model, namely sound intensity, is reduced, it is at least possible that a reduction in response might propagate itself along the chain from step 1 to step 3, resulting in a consequent reduction of the output signal, in this case the number of words chosen. In other words, that the reduced physiological arousal associated with the hypothesised lower volume of sound delivered to the autism group in the experiment, might lead to a reduction in the conscious awareness of cognitive emotions, and therefore both to reduced insight into internal emotional state and a reduction of the ability to respond to requests to describe those states. Therefore a failure to include the reduced sound volume as a covariate might, arguably, result in an appearance, but a spurious appearance, of impoverished emotional response to music.

On the other hand, favouring the straightforward acceptance of wordcount as a consequence of alexithymia is the strong correlation of one of the alexithymia factors (BVAQ-B factor 3: “poor insight”) with wordcount. The most obvious one is that it is difficult to see how this would explain alexithymia in real world situations, which is the aspect directly measured by the BVAQ-B. While it is conceivable that a voluntary reduction in sound volume is sufficient to explain the reduction in reported emotions in the autism group, there is no clear equivalent to this mechanism in normal life. It is implausible to explain the inability of people with autism to understand their own emotions, in terms of a self-imposed reduction in emotional response on their part through withdrawal from sensory stimuli. It may well be that the social anxieties often attendant on the condition will lead people on the autism spectrum to withdraw from social contact. It may be also that sensory hypersensitivity will compel some of these individuals to avoid excessively bright lights or harsh noises. But it is difficult to imagine any mechanism by which this behaviour could be responsible for their "mind blindness" towards the emotions of themselves or others.

6.3.10 Modifying the model to take account of the subjective/objective confound

We have, therefore, to seek another explanation and another mechanism. It appears that the initial model, connecting the physiological response directly to cognitive emotional response, is oversimplified. A connection between the subjective level of sound perception and cognitive emotional appraisal has therefore been tentatively suggested, which supplements or modifies step 2 in the model, and has been identified accordingly as Step 2'.

The data from the current series of experiments is not sufficient to determine further whether this model has credibility, and if so, how strong this additional factor may be. Future studies might explore this aspect. The obvious way to do so would be to conduct a comparative study of autism and control groups in which both sound levels and subjective measures of loudness were carefully measured, and input as IVs into a multiple regression with wordcount as the DV. This should enable the independent contributions of subjective loudness levels and actual sound volume to be disembedded from one another. Clearly, there would need to be careful controls on the subjectively perceived volume of sound. However, if participants in the autism group were given physical control over the loudness, and tasked with increasing volume to the "maximum level compatible with comfort", and then marking a further point on the subjective volume scale as the "most comfortable listening level", this would provide the basis for two points on the subjective scale that could be used for comparison with a control group.

6.3.11 Limitations due to restrictive word list and set of musical items

The word list used in the study in chapter 5 (Appendix 4.2) consisted mainly of words with positive affective connotations, with only a minority (test, sad, anxious, scared, wistful) carrying negative or unpleasant associations. This reflects the fact that the autism group comments from which these were derived were primarily describing the effects on them of music which they chose to listen to, which would therefore be positive; and the pilot group happened to consist of individuals of whom the majority clearly liked the sample musical items. There was therefore little scope for participants to indicate disapproval or dislike of the items, when it came to the main part of the experiment.

The variation in the "experience" variable in both the control and autism groups in chapter 5, as well as the existence of well defined "classical" and "pop" groups in the study reported in chapter 3, suggests that reactions to standard sets of musical items are highly likely to be conditioned in part by the personal musical tastes of the participant, and, equally, by the extent of their prior exposure to this type of music. The set of items used in the chapter 5 study was predominantly of a type that a fan of pop music might not find particularly appealing. The style of music might appear too "old-fashioned" to a fan of Gorillaz or Ke$ha, and the individual items varied greatly in intrinsic musical merit (one item was described in the pilot as sounding like "elevator music").

It seems likely that a large proportion of the unexplained variance in the analysis may be due to these facts, and future research along these lines should take this into account by providing a wider range of musical types, sufficient to include items appealing to a greater proportion of the participants. It would also be useful to include subjective assessments of liking of individual items by participants, to enable GSR reactions to be differentiated as between emotional reactions of strong like and strong dislike. The list of emotion words should also include options to indicate irritation, boredom and possibly other reactions induced by musical items aversive to participants.

6.3.12 Issues around the qualitative/quantitative divide

The use of qualitative research to illuminate new areas of study appears, in the author's view, to be undervalued in research. It can be thought of metaphorically as a way of making a rough sketch map of a new country: once the main mountains, rivers and so on have been marked in, it is then possible to measure the distances between them all and construct a proper chart. But it helps to have the sketch map first. In the same way, qualitative work can be invaluable in guiding later, more precise quantitative studies. The open-ended research question in the typical qualitative study (roughly translated as "What is the nature of [the phenomenon of interest]?") can then be refined into the category of research hypothesis that we are all familiar with, in the quantitative follow-up. So, in a simplified example, the outcome of a qualitative analysis might be to identify certain variables X and Y as being of major interest and importance, and the role of a subsequent quantitative analysis might be to determine whether, and to what extent, X correlates with Y.

Quantitative procedures invariably require the clear statement of a research hypothesis, which can be tested using numerical variables subject to a statistical significance analysis. However, in studying the area examined in chapter 3 of this thesis, the case has already been made in that chapter that insufficient prior data existed to enable a quantitative hypothesis to be put forward with any degree of plausibility. In these circumstances, the use of a qualitative method in chapter 3, followed by a quantitative study in chapters 4 and 5, appears both logically justifiable in theory and effective in practice.

One disadvantage of the basic qualitative design method used in chapter 3 is the lack of a comparison group. Whilst this is not uncommon in qualitative studies, it had the disadvantage that comparing my results with those in the literature on music and emotion in TD adults, required the assumption that those studies had been conducted in circumstances in which their results could be directly compared with those from my autism group. Given that the TD studies did not focus directly on the salient aspects of the chapter 3 study, this assumption may have been misplaced. The key result of chapter 3, that the sample used internal arousal language rather than externally focused emotion words, had not been directly addressed by previous TD studies, and therefore requires independent confirmation.

In advance of my own study it was impossible to predict this result, and therefore to deduce that a comparison group was in fact desirable. My attempt to make up for this shortcoming by building in a comparison between the two categories of words (internal and external) in the quantitative study, failed in its aim because the circumstances of the latter study were different, involving as they did a list of pre-existing words. Ideally, it would be advisable to repeat the study of chapter 3 with a group of TD adults, and to analyse specifically the words that they used, spontaneously and uncued, to describe their emotional reactions to music. Only then would it be possible to deduce with any degree of confidence that the internal/external distinction really exists.

6.3.13 Possible difficulties with gender balance

The existence of a gender imbalance in the diagnosed ASD population means that any control group which is gender-matched to a randomly sampled autism group is necessarily predominantly male. The only alternative to this would be to recruit an unusually large number of female participants in the autism group, and balance this with a gender balanced control group. However, given the low incidence of ASD in women, this would bring severe recruitment problems. The outcome of this situation is that the existence of any gender effect in autism studies is difficult to establish, because the power of any statistical test involving severely imbalanced group sizes across the variable of interest is known to be low.

This therefore restricts the conclusions that can be drawn as to gender effects in autism studies, particularly in the kind of experiment undertaken in this thesis, in which sample sizes are comparatively small. Furthermore, the study reported in chapter 2 on the empathic effects of music is subject to a substantial gender bias in the opposite direction, involving a preponderance of female participants. There is some evidence for an absence of gender effect on physiological arousal to music (Rickard, 2004) which suggests that such effects are unlikely to represent a threat to the validity of the results, but given that null effects are difficult to prove, this does not enable the issue to be shelved. For the time being, pending further studies involving larger numbers of gender-balanced participants, the question as to possible gender effects must remain open. It is, however, clear that such studies are warranted.

6.4 POSSIBLE CLINICAL APPLICATIONS[2]

6.4.1 Hypothesis as to the aetiology of alexithymia

It occurred to the author that the responsiveness to music in the population of high functioning adults on the autism spectrum that was demonstrated in chapter 3, coupled with the results of chapter 5 showing a specific impairment in the ability to name the

[2] This part of chapter 6 is a partially rewritten extract from Allen and Heaton (2010).

emotions evoked by music, might have applications to the possible treatment of the latter impairment. The reasoning was that the apparently complex problems surrounding the problem of type II (cognitive) alexithymia in autism presented themselves in a comparatively simple and accessible form, when studied in terms of the inability of people with autism to name the emotions evoked in them by music. So that a link existed, leading from a general inability (alexithymia) to a particular one (inability to identify emotions in music). It was speculated that the link might also operate in the reverse direction, so that teaching these individuals to overcome their inability to name and identify the emotions in music, might have an effect in reducing their alexithymia. In other words, it might be that a particular ability (to name emotions in music) might generalise to an ability to gain more general insights into one's own internal emotional life.

The chain of thought that began with this idea, was given further momentum by the readings of the literature on face-based empathic mechanisms cited in chapter 2, principally Goldman and Sripada (2005). These ideas led to the suggestions set out below. These suggestions go well beyond the empirical data presented in this thesis and are highly conjectural, but it is hoped to show that the conjectures are at least plausible, and represent the basis for a further programme of research.

The developmental trajectory of alexithymia in autism is not well understood. However, it is evident by definition that alexithymia reflects the *absence* of a typically observed behaviour rather than the *presence* of an atypical behaviour. A logical starting point for investigating it, would therefore be to pose the question not in terms of "why do certain people acquire alexithymia?", but rather as "how do typically developing individuals *avoid* alexithymia, i.e., how do they learn to identify and label their own internal states?" Conceptually, it is not obvious how humans arrive at a common language of terms that describe subjective and personal internal states. But it is appropriate to explore this question, since an understanding of the mechanism through which this ability is acquired in normal development, may provide clues as to how the mechanism breaks down in autism spectrum disorders, and how such a deficit may therefore be remediated.

I hypothesize that one principal means through which connections between the affective and cognitive domains may be formed in normal development is via the link between basic emotions and culturally universal (Ekman & Friesen, 1986), and possibly hard-wired, facial expressions. While a very young infant will not have acquired internal language labels for its emotion states, these are nevertheless reflected in its facial expressions. Research has shown that during typical mother-baby interactions, a mother will often mimic or mirror her baby's facial expressions (Marcelli, Tourrette, Kasolter-Pere, & Boinard, 2000). The typically developing infant attends to its mother's face and observes her expressions. It is hypothesized that repeated associations in the infant's developing brain between its mother's facial expressions and the infant's own internal feeling states, form the basis of its emotion-labelling ability.

It is suggested that this pre-linguistic association forms the scaffolding upon which the child subsequently constructs post-linguistic words and concepts. These are no longer purely internal and private (its own feeling) but are also external and public (a facial expression and a set of verbal labels). This interpersonal process may be supported culturally. For example, the widely popular 'Mr. Men' books include characters whose facial expressions are used to introduce a relatively wide repertoire of linked behaviours. 'Mr. Grumpy' shows a range of typically churlish behaviours, whereas 'Mr. Happy' smiles a lot and is pleasant to other people. The infant already knows what emotion goes with a 'Mr. Happy' face: it is the emotion that it was experiencing when it saw its mother with that same smiling facial expression (mirroring that of the infant). The book (or other cultural influence) teaches the infant the name, and other correlated qualities, of happiness.

Though it is plausible that associating an emotion with a facial expression does typically play a very significant role in typical development, it is not claimed that this mechanism is exclusive. It is likely that emotional expressiveness via the voice is also important, and for blind infants this channel, and perhaps also 'emotional' touch, being the only modalities available, would take on exclusive and dominant significance. Research in this area has not been extensive, but studies to date have indicated both the importance of touch in the development of exploratory behaviour in blind infants (Smitsman & Schellingerhout, 2000), and the therapeutic value of touch in stress reduction in premature infants (Whitley & Rich, 2008).

How would this explanation inform our understanding of alexithymia in autism? Research shows that social communicative cues are less salient for infants who are subsequently diagnosed with autism; for them, familiar faces – in particular those of their mothers – are not salient (Dawson et al., 2002). If the infant seldom orients to its mother's face, and does not observe her mimicry of its expressions, links between its own internal emotional states and her facial expressions will not be formed. It is suggested that the autistic infant's inability to benefit from such early bootstrapping experiences may, at least in part, account for the phenomenon of alexithymia in autism.

In developing a treatment for alexithymia in adulthood, one is therefore faced with the problem of finding a substitute for the developmental process that, in typically developing individuals, enables internal feeling states to be associated with external signs. Having missed the early opportunities to map the internal and external worlds onto one another, alternative methods are needed to remediate difficulties in adults with developmental alexithymia. One possibility might be to develop a system using a camera and computer to analyze the facial expressions of a person on the autism spectrum and feed back the results, effectively reconstructing the baby-mother mirroring system but in adulthood. A system along these lines is already being developed by Professor Rosalind Picard in the MIT Media Lab, and is being tested on college-aged ASD students with promising results (Picard, in press). Given that Picard's interesting work has not yet been fully developed and evaluated, it is worth attempting an approach that complements this work, and also builds on known strengths in musical appreciation in autism.

6.4.2 Systematic Induction of Emotion by Music

It was conjectured in chapter 5 above that the fundamental music emotions experienced by the autism group in that study might be similar to those experienced by typical listeners, and that group differences found in experimental results reflected alexithymia in the autism group. This conjecture suggested that it might be possible to develop a procedure based on the induction of specific feeling states in participants, using labelled items of music. The objective would be to use this effect to allow participants to learn an association between the emotions they experienced when listening to music, and the conventional names for those emotions. This would help participants to learn to distinguish their particular varieties of negative and positive emotions, by associating them with passages of music, so that when they experienced a feeling they could give it a music-related label. This would enable them to gain insights into their mental states, which would both reduce anxiety and enable them to function better socially.

It is possible that if this procedure were tried and proved effective, future approaches might combine Picard’s work with this musical approach. However, the approach that has been suggested differs from that of Picard in one fundamental way. In her work, moods and facial expressions are linked, and mood drives expression: the internal world drives the external one. By contrast, the approach suggested here would link internal and external worlds by using stimuli from the external world as the driver: it is proposed to use an external stimulus to induce a mood.

6.4.3 Associative Learning of Music/Emotion Links

Although the process that enables typical individuals to forge links between cognition and their internal feeling states is unitary, it can be divided into two components conceptually. When considering the case of autism it is fruitful to consider these two components separately. The first part of the link would consist of developing some cognitive hook or sign with which to associate a given feeling state. This sign might be part of a private language or it might, for example, in the case of synaesthesia, be a link with another modality. In this case a person might associate a particular mood with a shade of colour. This colour would be without associations or significance for anyone other than the synaesthete so the association would be both cognitive and private. The second link differs from the first in being culturally based, rather than intrapersonal. For example, when using the word “anger,” the individual can be confident that others from his/her culture will associate it with a specific mood state.

The point of this distinction is that an intervention that facilitates only the first part of the process might still be of value in mitigating the effects of alexithymia. This would be the case even if the individual did not develop the ability to fully identify his/her own emotions in culturally universal terms. As noted in chapter 1, Spinoza (as cited in Russell, 1961, p. 557) claimed that “an emotion which is a passion ceases to be a passion as soon as we form a clear and distinct idea of it.” He was concerned with the negative effect of “passions” and considered that it was worth aiming to cultivate the ability to form “clear and distinct ideas” of one’s internal feelings as a way of overcoming these negative effects. If Spinoza was right in this, then it may not matter if the “clear and distinct idea” cannot be formulated in terms of a common language. In the case of our hypothetical example, the individual might identify their internal state as a “lemon yellow” mood, or as “the way I always feel when I hear Brahms’s Piano Concerto.” While this may be of little help in enabling the individual to deal with the social implications of alexithymia, it may reduce the associated sense of helplessness and anxiety caused by the disorder. This is one instance where a private language may be comparable in usefulness with a public one.

Such an approach would use associative learning between a musical item and a mood as the basic mechanism. Given that there is no evidence that associative learning is impaired in autism (at least in high-functioning adults), this basis appears to be empirically sound. In order to make this approach work, it would be the aim to ensure that a given set of musical items could reliably induce a given set of moods in our participants. In previous piloting studies and in published work (e.g., Heaton, Allen, Williams et al., 2008), it has been shown that children with autism reliably matched extracts of classical music with visual depictions of mood states. This repertoire of music would be drawn upon in the proposed research, bearing in mind that systematic trialling of musical items might be necessary to establish the most effective repertoire for each participant: the ideal selection is likely to vary from one participant to another, if only because their musical tastes are certain to differ widely (see also, Allen et al., 2009).

Once it had been established that participants associated a particular musical extract with an external mood state, the salience of this association would be emphasised. Thus, participants would be taught to associate the musical cues with their own internal moods and to formulate verbal descriptions of these moods. This would probably involve a period of trial and error learning by participants. For example, take the case of one of the participants in the study reported in chapter 5, who recounted how she was unable to distinguish between her negative emotions: they all felt simply as if “the lights were going out” for her. For such a participant, the learning process would involve listening to a series of extracts until s/he found one that induced feelings similar to her current mood, and then labelling that mood by the corresponding item of music. It is envisaged that participants might be equipped with something like an iPhone with visible icons corresponding to a range of alternative musical items that they could play to themselves, in order to distinguish, on a trial and error basis, between the moods experienced on a daily or even hourly basis.

It should be emphasised that the approach outlined above is highly speculative. There is very little empirical evidence at present that suggests it would be effective, but it is proposed that it represents a research hypothesis meriting further investigation, given the pervasive nature of alexithymia in autism.

It is envisaged that in order to test this hypothesis, it would be necessary to carry out an independent group, pretest-posttest design, in which high functioning adults with autism would be randomly allocated to treatment and control groups. The pre-test and post-test phases would involve a comparison of data from multiple DVs: measures of alexithymia, depression and positive/negative affect. The purpose would be to measure whether, and to what extent, the treatment group improved on these measures over the course of the treatment period compared with the control group.

The treatment itself might involve an initial training period in which participants learned to associate a list of emotion words with a battery of musical items previously established as reliably evoking those emotions, followed by another session of reinforcement. At the end of this second session, they would be provided with an MP3 player onto which the musical extracts, suitably identified in the player by the appropriate emotion label, had been copied. The next phase, lasting perhaps one month, would involve participants in the treatment group keeping a record of occasions when they experienced a strong emotion which they could not identify. On such occasions they would rate the intensity, and valence (pleasant/unpleasant), of the emotion on a scale from 1 to 10, and by playing through the items on the MP3 player, attempt to identify the extract which matched the feeling most closely. They would also make a note of the circumstances in which the emotion occurred (to enable the researchers to deduce its probable nature). The analysis would require analysis of changes, if any, in alexithymia and affect over the treatment period of treatment compared with control groups. Qualitative analysis of the diary phase might reveal more detail about the relative effectiveness of the treatment between different emotion and mood states.

REFERENCES

Abecasis, D., Brochard, R., del Rio, D., Dufour, A., & Ortiz, T. (2009). Brain Lateralization of Metrical Accenting in Musicians. In S. DallaBella, N. Kraus, K. Overy, C. Pantev, J. S. Snyder, M. Tervaniemi, B. Tillmann & G. Schlaug (Eds.), *Neurosciences and Music III: Disorders and Plasticity* (Vol. 1169, pp. 74-78). Oxford: Blackwell Publishing.

Accordino, R., Comer, R., & Heller, W. B. (2007). Searching for music's potential: A critical examination of research on music therapy with individuals with autism. *Research in Autism Spectrum Disorders, 1*(1), 101-115.

Adams, A. S., Soumerai, S. B., Lomas, J., & Ross-Degnan, D. (1999). Evidence of self-report bias in assessing adherence to guidelines. *International Journal for Quality in Health Care, 11*(3), 187-192.

Allen, R. (in press). *Finding the words for it: how alexithymia can account for apparent deficits in the ability of an ASD group to describe their emotional responses to music*. Paper to be presented at the ICMPC11, August 2010.

Allen, R., & Heaton, P. (2010). Autism, music, and the therapeutic potential of music in alexithymia. *Music Perception, 27*(4), 251-261.

Allen, R., Hill, E., & Heaton, P. (2009). 'Hath charms to soothe ...' An exploratory study of how high-functioning adults with ASD experience music. *Autism, 13*(1), 21-41.

Allen, R., Hill, E., & Heaton, P. (2009). The Subjective Experience of Music in Autism Spectrum Disorder. In S. DallaBella, N. Kraus, K. Overy, C. Pantev, J. S. Snyder, M. Tervaniemi, B. Tillmann & G. Schlaug (Eds.), *Neurosciences and Music III: Disorders and Plasticity* (Vol. 1169, pp. 326-331). Oxford: Blackwell Publishing.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text Revision. ed.). Washington, DC: American Psychiatric Association.

Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994). The 20-Item Toronto-Alexithymia-Scale .1. Item Selection and Cross-Validation of the Factor Structure. *Journal of Psychosomatic Research, 38*(1), 23-32.

Bailey, P. E., & Henry, J. D. (2007). Alexithymia, somatization and negative affect in a community sample. *Psychiatry Research, 150*(1), 13-20.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a theory of mind. *Cognition, 21*(1), 37-46.

Baron-Cohen, S. (1989). The Autistic Child's Theory of Mind - a Case of Specific Developmental Delay. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 30*(2), 285-297.

Baron-Cohen, S. (1995). *Mindblindness: an essay on autism and theory of mind*. Cambridge, MA: MIT Press.

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences, 6*(6), 248-254.

Baron-Cohen, S. (2003). *The essential difference: men, women and the extreme male brain*. London, Allen Lane.

Baron-Cohen, S., Ring, H. A., Bullmore, E. T., Wheelwright, S., Ashwin, C., & Williams, S. C. R. (2000). The amygdala theory of autism. *Neuroscience and Biobehavioral Reviews, 24*(3), 355-364.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*(1), 5-17.

Batt-Rawden, K., & DeNora, T. (2005). Music and informal learning in everyday life. *Music Education Research, 7*(3), 289-304.

Bazeley, P. (2007). *Qualitative Data Analysis with NVivo*. London, Sage Publications Ltd.

Beaudet, A. L. (2007). Autism: highly heritable but not inherited. *Nature Medicine, 13*(5), 534-536.

Belmonte, M. K., Cook, E. H., Anderson, G. M., Rubenstein, J. L. R., Greenough, W. T., Beckel-Mitchener, A., et al. (2004). Autism as a disorder of neural information processing: directions for research and targets for therapy. *Molecular Psychiatry, 9*(7), 646-663.

Berthoz, S., & Hill, E. L. (2005). The validity of using self-reports to assess emotion regulation abilities in adults with autism spectrum disorder. *European Psychiatry, 20*(3), 291-298.

Bettelheim, B. (1967). *The Empty Fortress: Infantile Autism and the Birth of Self* New York, NY: The Free Press.

Bever, T. G., & Chiarello, R. J. (2009). Cerebral Dominance in Musicians and Nonmusicians (Reprinted). *Journal of Neuropsychiatry and Clinical Neurosciences, 21*(1), 94-96.

Bigand, E., Vieillard, S., Madurell, F., Marozeau, J., & Dacquet, A. (2005). Multidimensional scaling of emotional responses to music: The effect of musical expertise and of the duration of the excerpts. *Cognition & Emotion, 19*(8), 1113-1139.

Bird, G., Silani, G., Brindley, R., White, S., Frith, U., & Singer, T. (2010). Empathic brain responses in insula are modulated by levels of alexithymia but not autism. *Brain, 133*, 1515-1525.

Bishop, D. T., Karageorghis, C. I., & Kinrade, N. P. (2009). Effects of Musically-Induced Emotions on Choice Reaction Time Performance. *Sport Psychologist, 23*(1), 59-76.

Bonnel, A., Mottron, L., Peretz, I., Trudel, M., Gallun, E., & Bonnel, A. M. (2003). Enhanced pitch sensitivity in individuals with autism: A signal detection analysis. *Journal of Cognitive Neuroscience, 15*(2), 226-235.

Borg, E. (1977). Tail Artery Response to Sound in Unanesthetized Rat. *Acta Physiologica Scandinavica, 100*(2), 129-138.

Bormann-Kischkel, C., Vilsmeier, M., & Baude, B. (1995). The Development of Emotional Concepts in Autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 36*(7), 1243-1259.

Boso, M., Emanuele, E., Minazzi, V., Abbamonte, M., & Politi, P. (2007). Effect of long-term interactive music therapy on Behavior profile and musical skills in young adults with severe autism. *Journal of Alternative and Complementary Medicine, 13*(7), 709-712.

Bouhuys, A. L., Bloem, G. M., & Groothuis, T. G. G. (1995). Induction of Depressed and Elated Mood by Music Influences the Perception of Facial Emotional Expressions in Healthy-Subjects. *Journal of Affective Disorders, 33*(4), 215-226.

Bowler, D. (2007). *Autism Spectrum Disorders*. London: Wiley.

Bradley, M., & Lang, P. (2000). Measuring emotion: behavior, feeling, and physiology. In D. Lane & L. Nadel (Eds.), *Cognitive Neuroscience of Emotion*. Oxford: Oxford University Press.

Capps, L., Yirmiya, N., & Sigman, M. (1992). Understanding of Simple and Complex Emotions in Nonretarded-Children with Autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 33*(7), 1169-1182.

Caron, M. J., Mottron, L., Berthiaume, C., & Dawson, M. (2006). Cognitive mechanisms, specificity and neural underpinnings of visuospatial peaks in autism. *Brain, 129*, 1789-1802.

Christopher, A. N., Saliba, L., & Deadmarsh, E. J. (2009). Materialism and well-being: The mediating effect of locus of control. *Personality and Individual Differences, 46*(7), 682-686.

Courchesne, E., Townsend, J., Akshoomoff, N. A., Saitoh, O., et al. (1994). Impairment in Shifting Attention in Autistic and Cerebellar Patients. *Behavioral Neuroscience, 108*(5), 848-865.

Darwin, C. (1872). *The Expression of the Emotions in Man and Animals*. London: John Murray.

David, A. S. (1989). Perceptual Asymmetry for Happy Sad Chimeric Faces - Effects of Mood. *Neuropsychologia, 27*(10), 1289-1300.

Dalton, K. M., Nacewicz, B. M., Johnstone, T., Schaefer, H. S., Gernsbacher, M. A., Goldsmith, H. H., et al. (2005). Gaze fixation and the neural circuitry of face processing in autism. *Nature Neuroscience, 8*(4), 519-526.

Dawson, G., Carver, L., Meltzoff, A. N., Panagiotides, H., McPartland, J., & Webb, S. J. (2002). Neural correlates of face and object recognition in young children with autism spectrum disorder, developmental delay, and typical development. *Child Development, 73*(3), 700-717.

Deborde, A. S., Berthoz, S., Perdereau, F., Godart, N., Corcos, M., & Jeammet, P. (2004). Validity of the BVAQ : a study in eating disorder patients and controls.

Encephale-Revue De Psychiatrie Clinique Biologique Et Therapeutique, 30(5), 464-473.

Deborde, A. S., Berthoz, S., Wallier, J. M., Fermanian, J., Falissard, B., Jeammet, P., et al. (2008). The bermond-vorst Alexithymia questionnaire cutoff scores: A study in eating-disordered and control subjects. *Psychopathology, 41*(1), 43-49.

DeNora, T. (2002). Crossing the divide: Feminist perspectives on gender and music. VI. The role of music in intimate culture: A case study. *Feminism & Psychology, 12*(2), 176-181.

Deutsch, S. I., Urbano, M. R., Neumann, S. A., Burket, J. A., & Katz, E. (2010). Cholinergic abnormalities in autism: is there a rationale for selective nicotinic agonist interventions? *Clinical Neuropharmacology, 33*(3), 114-120.

Dunlap, W. P., & Kemery, E. R. (1987). Failure to Detect Moderating Effects - Is Multicollinearity the Problem. *Psychological Bulletin, 102*(3), 418-420.

Dunn, L., Dunn, L., Whetton, C., & Burley, J. (1997). *British picture vocabulary scale*. London: NFER-Nelson.

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect*. Palo Alto, California: Consulting Psychologists Press.

Ekman, P., & Friesen, W. V. (1986). A New Pan-Cultural Facial Expression of Emotion. *Motivation and Emotion, 10*(2), 159-168.

Fitzgerald, M., & Bellgrove, M. A. (2006). The overlap between alexithymia and Asperger's syndrome. *Journal of Autism and Developmental Disorders, 36*(4), 573-576.

Frith, U., & Happé, F. (1994). Autism - Beyond Theory of Mind. *Cognition, 50*(1-3), 115-132.

Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., et al. (2009). Universal Recognition of Three Basic Emotions in Music. *Current Biology, 19*(7), 573-576.

Gallagher, S. (2007). Logical and phenomenological arguments against simulation theory In D. Hutto & M. Ratcliffe (Eds.), *Folk Psychology Re-assessed*. Berlin: Springer.

Gibbs, G. R. (2002). *Qualitative Data Analysis: Exploration with NVivo*. Buckingham: Open University Press.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitiative research*. New York: Aldine.

Goldman, A. I., & Sripada, C. S. (2005). Simulationist models of face-based emotion recognition. *Cognition, 94*(3), 193-213.

Gomes, E., Pedroso, F. S., & Wagner, M. B. (2008). Auditory hypersensitivity in the autistic spectrum disorder. *Pro Fono, 20*(4), 279-284.

Happé, F. G. E. (1995). The Role of Age and Verbal-Ability in the Theory of Mind Task-Performance of Subjects with Autism. *Child Development, 66*(3), 843-855.

Happé, F. (1999). Autism: cognitive deficit or cognitive style? *Trends in Cognitive Sciences, 3*(6), 216-222.

Happé, F., Ronald, A., & Plomin, R. (2006). Time to give up on a single explanation for autism. *Nature Neuroscience, 9*(10), 1218-1220.

Haviland, M. G., Warren, W. L., & Riggs, M. L. (2000). An observer scale to measure alexithymia. *Psychosomatics, 41*(5), 385-392.

Heaton, P. (2003). Pitch memory, labelling and disembedding in autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 44*(4), 543-551.

Heaton, P. (2005). Interval and contour processing in autism. *Journal of Autism and Developmental Disorders, 35*(6), 787-793.

Heaton, P. (2009). Assessing musical skills in autistic children who are not savants. *Philosophical Transactions of the Royal Society B-Biological Sciences, 364*(1522), 1443-1447.

Heaton, P., & Allen, R. (2009). *"With concord of sweet sounds..." New perspectives on the diversity of musical experience in autism and other neurodevelopmental conditions*. Paper presented at the The Neurosciences and Music III, 2008.

Heaton, P., Allen, R., Williams, K., Cummins, O., & Happé, F. (2008). Do social and cognitive deficits curtail musical understanding? Evidence from autism and Down syndrome. *British Journal of Developmental Psychology, 26*, 171-182.

Heaton, P., Davis, R. E., & Happe, F. G. E. (2008). Research note: Exceptional absolute pitch perception for spoken words in an able adult with autism. *Neuropsychologia, 46*(7), 2095-2098.

Heaton, P., Hermelin, B., & Pring, L. (1998). Autism and pitch processing: A precursor for savant musical ability? *Music Perception, 15*(3), 291-305.

Heaton, P., Hermelin, B., & Pring, L. (1999). Can children with autistic spectrum disorders perceive affect in music? An experimental investigation. *Psychological Medicine, 29*(6), 1405-1410.

Heaton, P., Pring, L., & Hermelin, B. (1999). A pseudo-savant: A case of exceptional musical splinter skills. *Neurocase, 5*(6), 503-509.

Heaton, P., Pring, L., & Hermelin, B. (2001). Musical processing in high functioning children with autism. In *Biological Foundations of Music* (Vol. 930, pp. 443-444).

Heaton, P., & Wallace, G. L. (2004). Annotation: The savant syndrome. *Journal of Child Psychology and Psychiatry, 45*(5), 899-911.

Heaton, P., Williams, K., Cummins, O., & Happe, F. (2008). Autism and pitch processing splinter skills: A group and subgroup analysis. *Autism, 12*(2), 203-219.

Heaton, P., Williams, K., Cummins, O., & Happe, F. G. E. (2007). Beyond perception: Musical representation and on-line processing in autism. *Journal of Autism and Developmental Disorders, 37*(7), 1355-1360.

Hill, A. L. (1977). Idiot Savants - Rate of Incidence. *Perceptual and Motor Skills, 44*(1), 161-162.

Hill, E. L. (2004). Evaluating the theory of executive dysfunction in autism. *Developmental Review, 24*(2), 189-233.

Hill, E. L., & Berthoz, S. (2006). Response to "Letter to the editor: The overlap between Alexithymia and Asperger's syndrome", Fitzgerald and Bellgrove, Journal of Autism and Developmental Disorders, 36(4). *Journal of Autism and Developmental Disorders, 36*(8), 1143-1145.

Hill, E., Berthoz, S., & Frith, U. (2004). Brief report: Cognitive processing of own emotions in individuals with autistic spectrum disorder and in their relatives. *Journal of Autism and Developmental Disorders, 34*(2), 229-235.

Hill, E. L., Sally, D., & Frith, U. (2004). Does mentalising ability influence cooperative decision-making in a social dilemma? Introspective evidence from a study of adults with autism spectrum disorder. *Journal of Consciousness Studies, 11*(7-8), 144-161.

Hobson, R. P. (1993). *Autism and the development of mind*. Hove, Lawrence Erlbaum Associates.

Huron, D. (2001). Is music an evolutionary adaptation? *Biological Foundations of Music, 930*, 43-61.

Huron, D. (2006). *Sweet Anticipation: Music and the Psychology of Expectation*. Cambridge, Mass.: MIT Press.

Jarvinen-Pasley, A., & Heaton, P. (2007). Evidence for reduced domain-specificity in auditory processing in autism. *Developmental Science, 10*, 786-793.

Jarvinen-Pasley, A., Pasley, J., & Heaton, P. (2008). Is the linguistic content of speech less salient than its perceptual features in autism? *Journal of Autism and Developmental Disorders, 38*(2), 239-248.

Jarvinen-Pasley, A., Wallace, G. L., Ramus, F., Happe, F., & Heaton, P. (2008). Enhanced perceptual processing of speech in autism. *Developmental Science, 11*(1), 109-121.

Jolliffe, T., & Baron-Cohen, S. (1999). A test of central coherence theory: linguistic processing in high-functioning adults with autism or Asperger syndrome: is local coherence impaired? *Cognition, 71*(2), 149-185.

Juslin, P. N., & Vastfjall, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences, 31*(5), 559-575.

Just, M. A., Cherkassky, V. L., Keller, T. A., Kana, R. K., & Minshew, N. J. (2007). Functional and anatomical cortical underconnectivity in autism: Evidence from an fMRI study of an executive function task and corpus callosum morphometry. *Cerebral Cortex, 17*(4), 951-961.

Kaplan, R. S., & Steele, A. L. (2005). An analysis of music therapy program goals and outcomes for clients with diagnoses on the autism spectrum. *Journal of Music Therapy, 42*(1), 2-19.

Kern, P., Wolery, M., & Aldridge, D. (2007). Use of songs to promote independence in morning greeting routines for young children with autism. *Journal of Autism and Developmental Disorders, 37*(7), 1264-1271.

Khalfa, S., Bruneau, N., Roge, B., Georgieff, N., Veuillet, E., Adrien, J. L., et al. (2004). Increased perception of loudness in autism. *Hearing Research, 198*(1-2), 87-92.

Khalfa, S., Isabelle, P., Jean-Pierre, B., & Manon, R. (2002). Event-related skin conductance responses to musical emotions in humans. *Neuroscience Letters, 328*(2), 145-149.

Klin, A., Jones, W., Schultz, R., & Volkmar, F. (2003). The enactive mind, or from actions to cognition: lessons from autism. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 358*(1430), 345-360.

Klin, A., Volkmar, F. R., & Sparrow, S. S. (1992). Autistic Social Dysfunction - Some Limitations of the Theory of Mind Hypothesis. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 33*(5), 861-876.

Knickmeyer, R. C., Baron-Cohen, S., Auyeung, B., & Ashwin, E. (2008). How to test the Extreme Male Brain Theory of Autism in terms of foetal androgens? *Journal of Autism and Developmental Disorders, 38*(5), 995-996.

Konecni, V. J. (2003). Music and emotion: Theory and research. *Music Perception, 20*(3), 332-341.

Lathe, R. (2006). *Autism, brain and environment*. London: Jessica Kingsley

Lamont, A. (2003). Toddlers' musical preferences - Musical preference and musical memory in the early years. *Neurosciences and Music, 999*, 518-519.

Lamont, A. (2009). *University students' strong experience of music*. Paper presented at the International Conference on Music and Emotion, 2009.

Lane, R. D., Sechrest, L., & Riedel, R. (1998). Sociodemographic correlates of alexithymia. *Comprehensive Psychiatry, 39*(6), 377-385.

Lawrence, A. D., Calder, A. J., McGowan, S. V., & Grasby, P. M. (2002). Selective disruption of the recognition of facial expressions of anger. *Neuroreport, 13*(6), 881-884.

Le Doux, J. (1998). *The emotional brain : the mysterious underpinnings of emotional life* London: Psychology Press.

Lenzi, D., Trentini, C., Pantano, P., Macaluso, E., Iacoboni, M., Lenzi, G. L., et al. (2009). Neural Basis of Maternal Communication and Emotional Expression Processing during Infant Preverbal Stage. *Cerebral Cortex, 19*(5), 1124-1133.

Levinson, J. (2006). *Contemplating Art*. Oxford: Oxford University Press.

Levitin, D. (2006). *This Is Your Brain on Music*. New York: Dutton.

Levitin, D. J., Cole, K., Lincoln, A., & Bellugi, U. (2005). Aversion, awareness, and attraction: investigating claims of hyperacusis in the Williams syndrome phenotype. *Journal of Child Psychology and Psychiatry, 46*(5), 514-523.

Lewis, J., Elman, J., & Courchesne, E. (2005). Increased corpus callosum size in young children with autism. *Journal of Cognitive Neuroscience*, 194-194.

Lewis, M. D. (2005). Bridging emotion theory and neurobiology through dynamic systems modeling. *Behavioral and Brain Sciences, 28*(2), 169-245.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic Enquiry*. Beverly Hills, CA.: Sage Publications, Inc.

Lombardo, M. V., Barnes, J. L., Wheelwright, S. J., & Baron-Cohen, S. (2007). Self-Referential Cognition and Empathy in Autism. *Plos One, 2*(9), 11.

Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., et al. (1989). Autism Diagnostic Observation Schedule - a Standardized Observation of Communicative and Social-Behavior. *Journal of Autism and Developmental Disorders, 19*(2), 185-212.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., et al. (2000). The Autism Diagnostic Observation Schedule-Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30*(3), 205-223.

Losch, A. (2006). Combining quantitative methods and grounded theory for researching e-reverse auctions. *Libri, 56*(3), 133-144.

Marcelli, D., Tourrette, C., Kasolter-Pere, M. A., & Boinard, S. (2000). A depressed mother observes her child. A study of the search for behavioral signs which differentiate between depressed mothers and a control group of mothers in the neonatal period. *Psychiatrie De L Enfant, 43*(2), 541-586.

Meyer, L. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.

Mitchell, M. L., & Jolley, J. M. (2007). *Research Design Explained*. Belmont, California: Thomson Higher Education.

Mithen, S. (2009). The Music Instinct: The Evolutionary Basis of Musicality. In S. DallaBella, N. Kraus, K. Overy, C. Pantev, J. S. Snyder, M. Tervaniemi, B. Tillmann & G. Schlaug (Eds.), *Neurosciences and Music III: Disorders and Plasticity* (Vol. 1169, pp. 3-12). Oxford: Blackwell Publishing.

Montgomery, K. J., & Haxby, J. V. (2008). Mirror neuron system differentially activated by facial expressions and social hand gestures: A functional magnetic resonance imaging study. *Journal of Cognitive Neuroscience, 20*(10), 1866-1877.

Mosconi, M., Zwaigenbaum, L., & Piven, J. (2006). Structural MRI in autism: Findings and future directions. *Clinical Neuroscience Research, 6*(3-4), 135-144.

Mottron, L., & Burack, J. A. (2001). Enhanced perceptual functioning in the development of autism. In J. A. Burack, T. Charman, N. Yirmiya & P. R. Zelazo (Eds.), *The development of autism: perspectives from theory and research* (pp. 131-148). Mahwah, N.J.: Lawrence Erlbaum Associates.

Mottron, L., Peretz, I., Belleville, S., & Rouleau, N. (1999). Absolute pitch in autism: A case study. *Neurocase, 5*(6), 485-501.

Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., & Fried, I. (2010). Single-Neuron Responses in Humans during Execution and Observation of Actions. *Current Biology, 20*(8), 750-756.

Muller, M., Hofel, L., Brattico, E., & Jacobsen, T. (2009). Electrophysiological correlates of aesthetic music processing. *Annals of the New York Academy of Sciences, 1165*, 355-358.

Mustafa, Z. (2009). The Taliban and music [Electronic Version]. *Dawn*. Retrieved 22 January 2010 from http://www.dawn.com/wps/wcm/connect/dawn-content-library/dawn/news/pakistan/14-the-taliban-and-music-zj-03.

Nettelbeck, T., & Young, R. (1996). Intelligence and savant syndrome: Is the whole greater than the sum of the fragments? *Intelligence, 22*(1), 49-67.

Niedenthal, P. M., Brauer, M., Halberstadt, J. B., & Innes-Ker, A. H. (2001). When did her smile drop? Facial mimicry and the influences of emotional state on the detection of change in emotional expression. *Cognition & Emotion, 15*(6), 853-864.

Niedenthal, P. M., Halberstadt, J. B., Margolin, J., & Innes-Ker, A. H. (2000). Emotional state and the detection of change in facial expression of emotion. *European Journal of Social Psychology, 30*(2), 211-222.

North, A. C., Hargreaves, D. J., & Hargreaves, J. J. (2004). Uses of music in everyday life. *Music Perception, 22*(1), 41-77.

Patel, A. (2008). *Music, Language and the Brain*. New York, NY: Oxford University Press, Inc.

Pellicano, E. (2007). Links between theory of mind and executive function in young children with autism: Clues to developmental primacy. *Developmental Psychology, 43*(4), 974-990.

Peretz, I. (2001). Listen to the brain: a biological perspective on musical emotions. In P. Juslin & J. Sloboda (Eds.), *Music and Emotion*. Oxford: Oxford University Press.

Picard, R. W. (In press). Future affective technology for autism and emotion communication. *Philosophical Transactions of the Royal Society B: Biological Sciences*.

Pidgeon, N., & Henwood, K. (1997). Using grounded theory in psychological research. In N. Hayes (Ed.), *Doing Qualitative Analysis in Psychology* (pp. 245-273). Hove: Psychology Press.

Pignatiello, M. F., Camp, C. J., & Rasar, L. A. (1986). Musical Mood Induction - an Alternative to the Velten Technique. *Journal of Abnormal Psychology, 95*(3), 295-297.

Pinker, S. (1999). *How the Mind Works*. London: Penguin Books.

Preston, S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences, 25*(1), 1-72.

QSR International (2006). *NVivo 7, software data package.* Doncaster, Australia.

Quintin, E. M., Bhatara, A. K., Poissant, H., Fombonne, E., & Levitin, D. (2009). Emotion Recognition in music in high functioning adolescents with autism spectrum disorders. Manuscript submitted for publication.

Rickard, N. S. (2004). Intense emotional responses to music: a test of the physiological arousal hypothesis. *Psychology of Music, 32*(4), 371-388.

Rimland, B. (1978). Savant capabilities of autistic children and their cognitive implications. In G. Serban (Ed.), *Cognitive defects in the development of mental illness* (pp. 43-65). New York: Brunner/Mazel.

Rosenhall, U., Nordin, V., Sandstrom, M., Ahlsen, G., & Gillberg, C. (1999). Autism and hearing loss. *Journal of Autism and Developmental Disorders, 29*(5), 349-357.

Russell, B. (1961). *History of Western Philosophy*. London: Allen & Unwin.

Rutter, M. (2002). The interplay of nature, nurture, and developmental influences - The challenge ahead for mental health. *Archives of General Psychiatry, 59*(11), 996-1000.

Sacks, O. (1996). *An Anthropologist on Mars*. New York: Random House.

Salimpoor, V. N., Benovoy, M., Longo, G., Cooperstock, J. R., & Zatorre, R. J. (2009). The rewarding aspects of music listening are related to degree of emotional arousal. *PLoS One, 4*(10), e7487.

Sandia National Laboratories. (1997). Scientists Use Digital Paleontology to Produce Voice of Parasaurolophus Dinosaur Retrieved January 22, 2010, from http://www.sandia.gov/media/dinosaur.htm

Schachter, S., & Singer, J. E. (1962). Cognitive, social, and physiological determinants of emotional state. *Psycholgical Review, 69*, 379-399.

Schultz, R. T. (2005). Developmental deficits in social perception in autism: the role of the amygdala and fusiform face area. *International Journal of Developmental Neuroscience, 23*(2-3), 125-141.

Sifneos, P. E. (2000). Alexithymia, clinical issues, politics and crime. *Psychotherapy and Psychosomatics, 69*(3), 113-116.

Silani, G., Bird, G., Brindley, R., Singer, T., Frith, C., & Frith, U. (2008). Levels of emotional awareness and autism: An fMRI study. *Social Neuroscience, 3*(2), 97-112.

Sizoo, B. B., van den Brink, W., Gorissen-van Eenige, M., Koeter, M. W., van Wijngaarden-Cremers, P. J. M., & van der Gaag, R. J. (2009). Using the Autism-Spectrum Quotient to Discriminate Autism Spectrum Disorder from ADHD in Adult Patients With and Without Comorbid Substance Use Disorder. *Journal of Autism and Developmental Disorders, 39*(9), 1291-1297.

Smitsman, A. W., & Schellingerhout, R. (2000). Exploratory behavior in blind infants: How to improve touch? *Infant Behavior & Development, 23*(3-4), 485-511.

Spencer, E., Ritchie, J., Lewis, J., & Dillon, L. (2003). *Quality in qualitative evaluation: a framework for assessing research evidence*. Retrieved 13 May 2008. from http://www.gsr.gov.uk/downloads/evaluating_policy/qqe_rep.pdf.

Strack, F., Martin, L. L., & Stepper, S. (1988). Inhibiting and Facilitating Conditions of the Human Smile - a Nonobtrusive Test of the Facial Feedback Hypothesis. *Journal of Personality and Social Psychology, 54*(5), 768-777.

Stuart, G. S., & Grimes, D. A. (2009). Social desirability bias in family planning studies: a neglected problem. *Contraception, 80*(2), 108-112.

Tager-Flusberg, H., & Sullivan, K. (2000). A componential view of theory of mind: evidence from Williams syndrome. *Cognition, 76*(1), 59-89.

Thayer, R. E. (1978). Toward a psychological theory of multidimensional activation (arousal). *Motivation and Emotion, 2*(1), 1-34.

Treffert, D. A. (1988). The Idiot Savant - a Review of the Syndrome. *American Journal of Psychiatry, 145*(5), 563-572.

van de Mortel, T. F. (2008). Faking it: social desirability response bias in self-report research. *Australian Journal of Advanced Nursing, 25*(4), 40-48.

Velten, E. (1968). A Laboratory Task for Induction of Mood States. *Behaviour Research and Therapy, 6*(4), 473-482.

Vorst, H. C. M., & Bermond, B. (2001). Validity and reliability of the Bermond-Vorst Alexithymia Questionnaire. *Personality and Individual Differences, 30*(3), 413-434.

Waldman, A. (2001). A Nation Challenged: The Law; No TV, No Chess, No Kites: Taliban's Code, From A to Z [Electronic Version]. *New York Times*. Retrieved 22 January 2010 from http://www.nytimes.com/2001/11/22/world/a-nation-challenged-the-law-no-tv-no-chess-no-kites-taliban-s-code-from-a-to-z.html.

Welberg, L. (2010). Mirrors, mirrors, everywhere? *Nature Reviews Neuroscience, 11*(6), 1.

Whipple, J. (2004). Music in intervention for children and adolescents with autism: A meta-analysis. *Journal of Music Therapy, 41*(2), 90-106.

White, S., Hill, E., Happe, F., & Frith, U. (2009). Revisiting the Strange Stories: Revealing Mentalizing Impairments in Autism. *Child Development, 80*(4), 1097-1117.

Whitley, J. A., & Rich, B. L. (2008). A double-blind randomized controlled pilot trial examining the safety and efficacy of therapeutic touch in premature infants. *Advances in Neonatal Care, 8*(6), 315-333.

Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., & Rizzolatti, G. (2003). Both of us disgusted in My Insula: The common neural basis of seeing and feeling disgust. *Neuron, 40*(3), 655-664.

Wigram, T., & Gold, C. (2006). Music therapy in the assessment and treatment of autistic spectrum disorder: clinical application and research evidence. *Child Care Health and Development, 32*(5), 535-542.

Wilkinson, L. (1999). Statistical methods in psychology journals - Guidelines and explanations. *American Psychologist, 54*(8), 594-604.

Woodbury-Smith, M. R., Robinson, J., Wheelwright, S., & Baron-Cohen, S. (2005). Screening adults for Asperger syndrome using the AQ: A preliminary study of its diagnostic validity in clinical practice. *Journal of Autism and Developmental Disorders, 35*(3), 331-335.

Young, R. L., & Nettelbeck, T. (1995). The Abilities of a Musical Savant and His Family. *Journal of Autism and Developmental Disorders, 25*(3), 231-248.

Zangwill, N. (2004). Against emotion: Hanslick was right about music. *British Journal of Aesthetics, 44*(1), 29-43.

Zwaigenbaum, L., Bryson, S., Rogers, T., Roberts, W., Brian, J., & Szatmari, P. (2005). Behavioral manifestations of autism in the first year of life. *International Journal of Developmental Neuroscience, 23*(2-3), 143-152.

APPENDIX 2.1: Musical Experience Questionnaire

1a) Have you had any formal musical training (not just general music classes at school)?

Please indicate below the extent of this.

(none)1…………2……………3………….4…………..5…………..6…………7(professional)

b) If yes, for how many years?

0……..1………2………..3………..4…………5………….6…………..7……………8

2) As a child, did you engage in any of the following musical activities for at least a year?

If so for how long per week?

a)	Individual music lesson	½ hr	1hr	1½hr	2+hrs
b)	Class music lessons	½ hr	1hr	1½hr	2+hrs
c)	Music therapy	½ hr	1hr	1½hr	2+hrs
d)	Dance/movement classes	½ hr	1hr	1½hr	2+hrs

3) How often did you, as a child, *choose* to listen to music at home?

Rarely

Moderately

Frequently: once per week…….2-4 times per week……..every day

4) How would you rate your *reaction* to music then?

e.g., music played on the radio or music centre, or music in films

Dislike Indifference Mild liking Moderately enthusiastic Very enthusiastic

5) At what age (if any) did you start getting personally interested in music?

6) How often do you choose to listen to music *now?*

Rarely

Moderately

Frequently: once per week…….2-4 times per week……..every day

7) What sort of music (if any) do you enjoy listening to now?

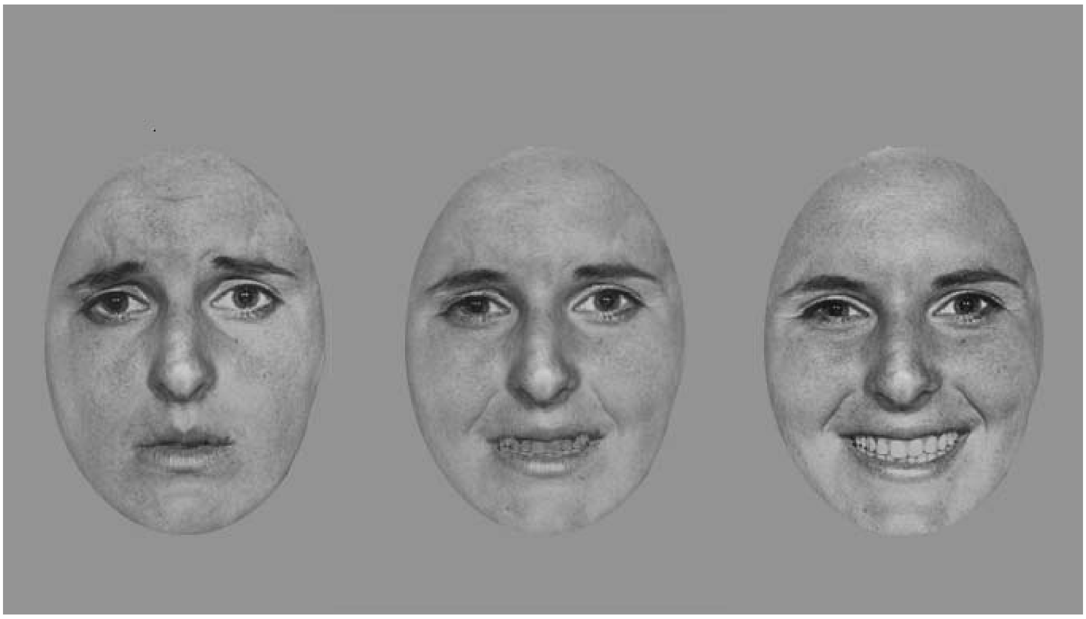
8) Do you regularly play a musical instrument, sing, or dance, now? (*Please tick*)

Yes □

No □

If so please give details:

APPENDIX 2.2: Sad, neutral and happy face exemplars



APPENDIX 3.1: Semi-structured questionnaire

1. Autobiographical memories about reactions to music:

Can you remember when you first began to get interested in music?

How old where you then?

What do you remember liking about it?

Can you think back to when you were really young, and you began to listen to music and to people talking? Did you prefer one to the other?

Do you think you might have liked music at this time? If you did, what did you like about it? Did it seem to make sense to you? If so, why?

How did your liking for music [if any] develop over time as you got older?

2. Current reactions to music:

What is it about music that makes you want to listen to it now – what does it do for you?

Do you see images when you listen to music: landscapes, shapes, people, colours?

How important to you are the lyrics in songs, as opposed to the music?

Does music make you feel emotional? For example, happy or sad or frightened or angry or anything like that?

What do you think it is it about the music that makes you feel like that? (If they have outlined 3 different emotions go through them separately).

Do you find different kinds of music, for instance classical music or jazz or pop music, makes you feel these things more than another?

What do you think it is about jazz (classical etc) that makes you feel like that?

If you watch a movie that is supposed to be scary, would it be less scary if you watched it without the music?

What about if it was supposed to be happy: would it seem less happy without the music?

Does the music ever help you understand what is going on when you watch films or TV?
If you were feeling really happy one day and you heard some sad music, would it change your mood? And what about if you were feeling sad and then heard happy music?

Do you have any particular piece of music that you associate with something happy that happened in the past? Or maybe something sad or exciting that happened in the past?

Do you have any particular piece of music that reminds you of a particular person?
If yes, why?

APPENDIX 3.2: Interview with participant GK

When did you first get interested in music?

At about 9 years, I first heard Oasis. But at about 11, 12 years, I don't know how I managed to though, my Mother died when I was 11, and I was feeling all kinds of feelings …

What sort of music did you like first?

It's quite embarrassing to talk about, but I remember when the Spice Girls came out, and they started a revolution with girl bands, in popular music the charts have been dominated by boy bands, I liked the Spice Girls, I didn't like their music so much but I needed female attention, really, it's silly now but I can look back and say the Spice Girls have led a revolution in popular music. Now there are lots of bands like that. Their lyrics are more infectious.

So you liked the idea of girl bands?

I think it was my young hormones coming into action too soon.

Can you take me through how your tastes in music evolved?

I was a bit of a lonely person when I was younger, not just because my Mum died and my Dad wasn't always around for me, also I found it hard to make friends and socialise, which is part of my condition, it's not so hard to socialise now, but I had that emptiness in me, and I was going through those growing up stages, so I listened to the radio and I listened to Heart 106.2 which played all the smooth love songs and kind of sad songs, and it reflected how I felt at the time, and later on I started to peg my interest to charts and started watching TOTP and I remember when Robbie Williams came out and wasn't doing very well until he released that song "Angels" in December '97, and all of a sudden everyone started liking him again, and I only knew he was from that boy band, and then I heard "Angels".

He quit the band and they didn't last long. I guess that song Angels helps people in a lot of ways either because it could be perceived as a love song, or a religious song, or an arm around your shoulder, or a shoulder to cry on, that kind of thing. So then he started to do really well and talking about growing up and stuff and I started listening to Capital FM and it was more mainstream, more variety with the Pepsi Chart there, and as I was getting older, I had just turned 13, I started to feel a bit better about myself and things generally, I had a good first year at school, and the second year didn't start off very well but then I began to fool around a bit and accept who I was, and I don't know how people perceived me, but I guess in my own world I was feeling good about myself, and I was very into the charts then, and would go out to the shops, and buy my favourite chart singles that were out, and every time I bought one I used to think, I'm helping them to get a chart position, I'm helping them to get to number one, I may be only buying one copy but I'm doing my bit, if 500,000 other people across the whole country are doing the same as me then the record will get to number one, and so I started thinking about how many weeks a song has been in the charts and what kind of songs…

There was another sad song that had been in the charts a while in '98, it was called "How do I live" by Leanne Rines, that was quite a nice song, and so I really got into pop stuff and that, and later on I was listening to Capital FM and sometimes people would do this thing called "Hall of Fame" where listeners could phone in and ask the DJ to pick three of their favourite songs of all time and if they're played they win a prize ... some of the songs they played were REM songs, and I thought I like this band, and then when I was in year 9, my art teacher said he liked REM so he let me borrow some of his REM CDs, and I had heard an REM song and went through every track on his CDs trying to find it … he let me record a couple of them onto cassette, and one of them was called "Automatic for the People", and I like it a lot, and .. this was like '98, '99, and things were not always going well, so after a good period in my life I was having a bad period, and I felt really down and out about myself, and sometimes I would listen to "Automatic for the People" and "Out of Time" as well, and when I listened to Automatic, it made me feel better about things, it was like a silver lining, and that was how I got to like REM so much.

When you say you felt better when you played their music, can you describe how that might have worked for you?

It reflected the way I was feeling, and it wasn't just one song, it was lots of songs, and I thought it's not just the way these songs are, it's the order in which they've been put on the CD, if it was in a different order it wouldn't sound the same, it wouldn't have the same feel … I was feeling a bit gloomy, not depressed, just sad that I was being bullied a lot, and I couldn't do anything much about it, and the mood from REM helped me to … it was as if there was a message from them, saying "don't worry about things, we'll play this music and I hope it makes you feel better".

So you thought it was a message for you from somebody who understood what you were going through?

Absolutely, and there must be hundreds of thousands of people across the world who have listened to that album and been touched by it …

Someone might say that if you are sad you would want to listen to happy music, but maybe they might also want to listen to music showing understanding of their sadness?

Yes, but also I realise now talking about this that there must have been a lot of people feeling a lot worse than I was, people on drugs or alcohol or feeling suicidal, I wasn't suicidal, just a bit down because of the way things had turned out for me.

Did other groups do this for you, like Radiohead or other groups?

I listen to Cold Play a lot now, I've got their albums, and … are we talking about now or in the past?

Take me through the music you liked after year 9 first. Did you move on from REM?

Not straight away, I just listened to whatever was around, most of them were good songs, and I enjoyed listening to them.

What did they do for you - was it always the same as REM?

Obviously being younger then I was a lot more energetic, and at that point I hadn't been taken abroad by alcohol, and then I smoked for a few years - I don't now, but I realised the difference between how I felt before I started smoking and how I feel since, there have been times since when I really felt down about stuff, what smoking has done to me has made me feel worse, kind of rotted away my insides, and with it my youthfulness, but yeah, so what I am trying to say is that when a song came out then I would be more able to move to it …

Did you go dancing?

If I was at parties I would a bit and I wouldn't care if people were looking at me funny or laughing I was just enjoying myself really …

What about your tastes now?

I've got an IPOD Nano, it's really useful, sometimes I just play different music depending on what I feel like I'm into, sometimes I declare myself to have an "album of the month" that I listen to, I'll play it every day on that month, until the end of the month, just to get a feel for it, if I haven't listened to it much before. I've got a 4 gigabyte IPOD, it's got about 850 songs on it at the moment, some of them I don't listen to because I'm a bit let down by some of them, I listen to mainly rock and indie and sometimes a bit of pop as well.

When I say "pop", there's two different kinds of pop, there's the trashy commercial stuff which you get with say Atomic Kitten and Take That and S Club Seven which I admit I did like a bit when I was younger, but I think everybody goes through that stage, when they say "I liked that cheesy band when I was younger", so I'm not ashamed to admit that. But there's that kind of pop, people say it has a "shelf life", I didn't realise what people meant by that until recently, it means it's not guaranteed to last for ever. But you see bands like the Beatles, and the Stones, and Simon and Garfunkel, they've been round for what seems like an eternity, at least the surviving members, but a shelf life band is lucky if it gets five years or … I've worked out what "shelf life" means, it *means say you go into a supermarket, you see something you want to buy to eat, and it's got a use by date on it, you have to take it by a certain time, I've figured out what that means, so …*

You've got those kinds of pop artists, and then you've got the more sophisticated kinds of pop, which I associate really with Elton John, Madonna, Kylie Minogue, Robbie Williams those kind of people, who do their own stuff, they don't get bossed around by people. I always associated Robbie Williams with rock, but now he does some rock tracks, but he's never going to be a rock star, he's just going to be up in the mainstream pop, but he has his own unique style, he doesn't really have a particular sound, which is a good thing, because if you sounded the same all the way through people would lose interest. You've got to be versatile.

Do you like the Red Hot Chili Peppers?

Yes. They have a fingerprint, but they don't always do the same thing.It's like Queen as well, I like Queen too, because a lot of people have tried to base their style on Queen, because it's rock but it's not hard core rock like Black Sabbath. It was just a few guys before, then Freddie Mercury came along, and he's into opera and ballet, and he joined up with Brian May and … they all liked to write their own things, Brian May was more into the guitar-themed songs, anthems, and Freddie Mercury would try to make a soap opera out of each song. What's good about Queen is that they've done so many great songs that people can think of, and sing, or they know the lyrics to (Bohemian Rhapsody?) *yes that did it from the start, there are a lot of people who like that song, and I've got the sheet music for it at home, and it's so complex, and Freddie Mercury did that by himself, no-one else just him, it's like - how do you write something like that, it's just ingenious, it's like no other song that's ever been played before, and that's why people love it so much, that's what got Queen famous … I haven't played that song for a while, I might play it a bit later, but … if you ask people in research to name as many Queen's songs as you can, they could at least come up with five or six.*

And their first greatest hits album, what's so great about it is, there are 17 tracks, and I noticed that because I'm into the history of bands, and statistics, all the songs that are on that album, all of those were top 20 songs, only one number one which was

Bohemian Rhapsody, but it was number one for nine weeks, and Queen did release other songs in seven years, because that's the period it covers, but some songs didn't make the top 20, and the ones that didn't, they didn't get onto that greatest hits album, so … I think it is one of the biggest selling albums of all time, only Michael Jackson has got more than they have.

At the time it was the Sergeant Pepper album that did that as well, whenever it came out, in '67 I think?

No it was '68. But what I mean is, the Chili Peppers always like to put on a good show for their crowd, and they're always loud and lively, and in your face, because none of them likes a boring rocker ..

Are you saying that you like something a bit loud sometimes?

Yes, I'm a lot of the time a restless character, and I got a bit carried away with Queen, because I wanted to perform, to perform on stage and have people adoring me, I just wanted attention, I wanted people to say "oh look at me", I'm a showman, which is like what Robbie Williams does at the moment, he does it really well, he's similar to Queen, not style-wise but he's versatile, he can do so many different things. He issued a greatest hits album and the songs on that one, they covered a seven year period as well, it just goes to show, you don't have to be around for ages to release a "greatest hits".

Do you ever listen to Enya?

No, I don't mean this in a sexist way but I'm not into girls' music so much, girl rock stars, girl guitar players like Katie Melluo and Katie Tunstall, it's not my scene, I just want to focus and see what guys have to say …

Do you like Radiohead?

I haven't listened to Radiohead much, I don't know what to make of them, I know they've done a couple of great all time albums, I like listening to the great ones; in Q magazine readers get to vote what are the great albums of all time, and I remember the *last time they did it, Radiohead had two of their albums in the top four, which was The Bends and OK Computer, and OK Computer had been number one the previous time but it was knocked off by Nirvana, do you know the album with the baby on the front? When someone's released a classic album and you listen to it, I think "I'm part of something great now", it's like when I listen to Automatic for the People or A Rush of Blood to the Head by Cold Play or By the Way by the Chili Peppers it makes me think these are great albums, I'm so happy I've listened to that, it's an all time classic, and I appreciate why it is as well.*

Can I try to summarise, you said you liked listening to REM's sad music because it matched your mood, whereas the Chili Peppers is very energetic music, would it be right to say in better moods you like listening to more arousing music?

I guess that's pretty close.

Do you ever have times when you're sad, if I listen to happy music it'll make you feel happy by association? Does that work for you?

No, I wish it did, but I'm not one of those people who say I'm feeling sad and I'm going to listen to something to change my mood, I'm not like that, music can't change my mood, only I *can change my mood.*

So you want something that is consistent with your mood, and enable you to bring it to the surface, would that be right?

Yes

So if you can bring it out, you will feel better about it?

I think so yes, because if I'm trying to relax I'm not going to listen to something that's loud or fast paced or … I'm going to chill out or something … Have you heard the Arctic Monkeys? I've got their first album called "Whatever people say" and I think that's an awesome album, it will be called a classic album one day, and I think it's awesome because it's very indie rock, and it's very working class as well, the lyrics are *like - we're students from Sheffield and we like to go out and have a good time - and the Killers, Hot Fuss: that's a great album, a classic album for the future … and I like, in contrast, I like .. one of my favourite albums at the present is Jack Johnson, "In Between Dreams", that's a very relaxing .. that's one I like to relax to.*

So when you want to relax, you listen to relaxing music and it makes you more relaxed?

Yes, but it's not relaxing in a sad way, it's relaxing in a sunny Sunday morning kind of relaxing.

What about listening to music that makes you the opposite of relaxed, ready for action, do you ever do that?

It's funny you say that, because the last three I was talking about, Arctic Monkeys, Jack Johnson and the Killers, I've got those three CD's in my CD changer in my car, I used to connect my IPOD to it, but the wire's broken, so if I'm out driving on the motorway I like to play the Killers to keep me in the driving mood, but if I'm like driving into London this morning for example, or anywhere urban where there's a lot of traffic and it's start-stop, I listen to something relaxing like REM or Jack Johnson, something a bit smooth so that I don't get so over-anxious and over-heated up, because there are a lot of idiot drivers out there who get you wound up so easily.

Getting off the track we've been following, I have a few specific questions. When you listen to music, do you ever see images, do pictures come into your mind like shapes, landscapes, people …

Oh, all the time, people mainly, people who made me feel awkward in the past, and stupid, silly things in my head which I don't like. If a part of a song comes on which is an instrumental bit and it makes me think of something stupid, I either turn the sound down or fast forward it. It depends what kind of song it is … if people have made me upset or angry in the past.

What is it about the song that prompts that association with people?

Well … I … I don't know. I've got this thing about people's surnames, it's like there are certain people's surnames I like, and that I don't like because of what their surnames begin with. It's not so much I don't like the person, but it's their surname. One of them I didn't like was the letter J, but that doesn't mean I'm not going to listen to Jack Johnson, because I like Jack Johnson .. there's the H as well .. which are related to these problems. That doesn't mean I didn't like your supervisor [surname Hill], *it doesn't work like that. Someone with a J or an H could be a nice person, or someone with a C, K or M which are my favourite surnames could be a really nasty person.*

Sometimes when I'm listening to Jack Johnson it makes me think of people whose surnames begin with the letter H, which is associated with J, and if I'm listening to someone whose surname begins with H, like The Darkness, it makes me think of people whose surnames begin with the letter J, I've got a crossover, but a lot of these people haven't been nice to me anyway, or they've upset me before, but I don't always think about people who've upset me, I think of good things as well, I think about silly things …

But on a more serious note, there are songs that - I'm sure you will be glad to hear this as part of your research - that make me feel sad, because I think the two things you're trying to find out about here are music and emotion, and last summer I had this quick - I won't say affair - association, with a girl, and I thought she liked me but, it didn't last long because we lived too far apart from each other, and she was a bit funny anyway, but I really liked her because I thought that she really liked me, and then Cold Play had released a new album by then, and it was played on the radio at the same time when all this was going on, and it …

Sometimes when I play that song and it makes me think about her, and in fact when I listen to the album ("X and Y") it makes be think about her a bit because she played it as well, some of the songs later in that album made me think about her as well emotionally, I felt loved up - not sex, I mean love - and Cold Play's second album that I mentioned before, "Rush of Blood to the Head", I think there was a couple of songs on there which not so much makes me think about her, but it makes me think of the countryside, because the countryside is nice and green and peaceful and you've got blue *sky and rivers and fresh air, and one of the songs was called "Warning Sign", this was from the second album, not the X and Y album, and it's not so much the lyrics that makes me think of her and the countryside, it's the mood of the song, the tempo of the song, and that's what I like about it, I don't … not that the lyrics make me think it's a sad song, for once I'm actually listening to the music, and not the words, because usually I listen to the words, but you see, you can like a song … you can either listen to the music, you can listen to the words, or you can listen to both, and you can think a song is nice, because even if the lyrics are telling a sad story, the music can make you think of something you like, and in this case trees and countryside and fresh air.*

Might it be that the music could be affecting you when you're just listening to the words?

You mean you're not giving 100% attention to the music, but it's filtering into your brain, and it just goes down into your body, and your soul and that. Also I think certain bands, I find easier to listen to and appreciate because of the surname thing, I liked Queen anyway, but listening to them is easier because of Mercury and Brian May and the same with Cold Play, with Chris Martin and Will Champion are two of the members, it doesn't matter that the other two members' surnames begin with B,

B is neutral?

No, B was the first one that got me into it, B is a bad one, but it doesn't … I try not to think of the bad ones so much, I mean why think of the bad ones when you can think of the good ones, so that helps as well a bit, and the Chili Peppers as well, Anthony Kiedis, and …

I can see why K is a good letter! [The participant's surname begins with K].

It wasn't a self bias, it was just the way it worked.

One of the things I'm interested in is the way music can affect people without their being aware of it, and one of the ways that can happen is if you're listening to a programme, or an opera, or a film, background music can change your mood, like in a horror movie … are you ever aware that the music is doing a job on your head?

Yes, that's what it's supposed to do, because it's like you're watching it as though you're watching a fictional but real life situation, but obviously if you were watching it for real, and you were there in the flesh you wouldn't hear the music, like in Rocky you get all the different music there like when he's training or he's in the fight itself, there's quite a few of them, I've watched all of them, you think the music, it really adds to the suspense, the feeling, the emotion, what is going on here, and it actually catches the moment really well, but if you were at a real life sporting event you wouldn't be hearing that music, and you wouldn't know what's going through the guys' heads, so you're in the comfort zone when you're watching it on the screen, but if you close your eyes, it's just like people playing the violins or whatever … I think I captured that really well!

Do you find that you find it difficult to work out what is happening in a film?

Sometimes I have to watch it over and over again to work out what is going on, but I never thought about it before, but I guess I don't always understand it, I can't really see into people's minds or what they're feeling, that doesn't mean I can't sympathise with people, in fact I always try and let the other person say what they have to say before I do, if this is real life I mean.

So in real life do you find it difficult to understand why people behave to you as they do?

Sometimes I feel that, but sometimes I see what the situation is, I think someone's having a bad day, or they could be late, or someone could have really upset them before, but if they just have a short threshhold or something, I should never take it personally, but some people, they take a lot of things out on me for no reason, but I say well if you're going to be like that to me, I just won't ever bother talking to you, I don't want to be treated like this, I don't see what I've done to deserve to be treated like this.

APPENDIX 3.3: Interview with participant WB

Can you remember when you first began to get interested in music? How old where you then?

Started doing piano at 7 or 8, but didn't really get into it until a teenager.

What do you remember liking about it?

Loved singing - was in church choir - learned to dance, possibly also ballet. Loved anything to do with music and movement.

Can you think back to when you were really young, and you began to listen to music and to people talking? Did you prefer one to the other?

I didn't talk – I was just happy to listen. I spent most of the time on my own. I used to go into a nearby wood and sing at the top of my voice, which was exhilarating. I had no interest in other children. I was on my own till age 4, when a brother was born, who gratified my parents by being normal and liking other kids.

Can you remember the time before you began to understand what people were saying to you? Do you have any memories about this?

No, I spoke early - and learned to read early as well, before going to school.

Can you remember when you first learned to speak? Do you have any memories about this, and about what it was like for you then?

No.

What is it about music that makes you want to listen to it now?

Familiarity - hence I like Classic FM, rather than radio 3 which tends to be unfamiliar. I like harmony more than melody. I can remember a passage from the Pearl Fishers, a *duet, and can hear both voices singing at once. The rules of harmony are mathematical.*

Does music make you feel emotional? For example, happy or sad or frightened or angry or anything like that?

I am either invigorated or left cold. I feel happiness when singing, but when listening, just contentment. Music is calming if I am anxious. It distracts me from unpleasant thoughts in my head. Listening to voices can also be calming, if there is no requirement to respond - eg listening to the radio. I have an alarm/radio, but the voices send me back to sleep now that I am retired and don't have to get up to go to work. I don't feel fear, anger, sadness or anxiety when listening to music.

Do you find different kinds of music, for instance classical music or jazz or pop music, makes you feel these things more than another?

I like trad jazz and the big band sound

What do you think it is about jazz (classical etc) that makes you feel like that?

Partly because it is associated with ballroom dancing, which I like. I have the impulse to move when I hear music, and beat time.

If you watch a movie that is supposed to be scary, would it be less scary if you watched it without the music?

I am not conscious of it. Scary music maybe makes me think something is happening.

What about if it was supposed to be happy: would it seem less happy without the music?

Yes.

Does the music ever help you understand what is going on when you watch films or TV?

Probably I understand social situations quite well, better than some normal people. But that is because of parental training: they would always tell me what was the right or wrong thing to say in situations.

If you were feeling really happy one day and you heard some sad music, would it change your mood? And what about if you were feeling sad and then heard happy music?

No.

Do you have any particular piece of music that you associate with something happy that happened in the past? Or maybe something sad or exciting that happened in the past?

Before my daughter was killed, she bought me some music: "Only You" by Platter - it was intended as a birthday present, and found in her possessions after her death. That would make me sad.

APPENDIX 3.4: Definitions of nodes and subnodes, and representative quotations from them

Definitions of nodes and sub-nodes

1. Motivation: reasons, aims or motivations which led participants to engage in the musical experience.

Subordinate Nodes:

1.1 Achievement: pleasure gained from increasing one's passive knowledge of music or one's technical ability in some aspect of composition or performance.

1.2 Aesthetic: feelings of pleasure that are derived from the conscious or intellectual perception of certain qualities of the music (eg beauty), rather than being a spontaneous emotional reaction to it.

1.3 Belonging: the pleasurable feeling induced by music in enabling listeners or performers to feel part of a wider community associated with that kind for music.

1.4 Enjoyment: simple pleasure in the process of making or listening to music, where this pleasure is not qualified by some other more complex motive than 'enjoyment'.

1.5 Mood Altering: the deliberate choice of listening to music in order to change mood state and attain a more desirable one, on a daily (short term) basis.

- 1.5.1 Valency Change: inducing moods of happiness or sadness with music
  - 1.5.1.1 Uses It
  - 1.5.1.2 Doesn't Use It
- 1.5.2 Arousal and Relaxation: inducing moods of arousal or relaxation with music
  - 1.5.2.1 Arousal: uses music to induce pleasurable arousal
  - 1.5.2.2 Relaxation: uses music to induce relaxation

1.6 Mood Congruent: the use of music to overcome negative feelings by listening to music of similar mood; also (less often) to the use of positive music to maintain a positive mood state.

1.7 Movement: the pleasure from the sense of physical movement induced by music, both literal and metaphorical (as in the phrase 'taken out of oneself').

1.8 Performance: refers to both actual performance of instrumental music (or an associated activity such as singing or dancing to music) and to the pleasure from imaginary performance while listening to music.

1.9 Therapeutic: listening deliberately to music in order to relieve a prolonged period of depression or unhappiness.

2. Characteristics: those aspects of the music that participants chose to play or listen to, that they perceived as significant reasons for choosing that particular music.

Subordinate nodes:

2.1 Appropriate Association: music felt to be appropriate to the context, or that induced pleasant thoughts by association, or film music that was helpful in establishing atmosphere by inducing the appropriate feelings.

2.2 Emotional Expressiveness: music whose attractive characteristics were described using typical emotional terms such as happy, sad, angry.
2.3 Energizing and Relaxing: the two subcategories have the obvious meaning here.
2.3.1 Relaxing or Slow.
2.3.2 Power, Speed, Beat.
2.4 Familiarity.
2.4.1 Familiarity a Requirement: music for which its familiarity to the listener made it pleasurable.
2.4.2 Variety a Requirement: music that was liked for its novelty or unfamiliarity.
2.5 Melody and Harmony: describing music that was liked for its melodic or harmonic qualities as distinct from other musical characteristics.
2.6 Message: music which was liked because the lyrics appeared meaningful or helpful to or directed at the listener, or where the philosophy or background of the group or band was understood and liked by the listener.

2.7 Social Connections: music which the listener liked, and to which they were introduced by a friend, or which had other associations of a social nature (eg the favourite tune of a significant other).
2.8 Structure and Pattern: music which was appreciated for its internal structure and pattern, or for the particular sequence of tunes on a CD, rather than for other qualities.
2.9 Technical Competence: the clarity of playing and quality of musicianship of the artists.
2.10 Timbre: music where the quality of the timbre (usually depth and richness of timbre) made the music likeable.

Direct quotations illustrating the nodes (each quotation under a particular subnode is from a different individual)
1.1 (achievement):
"I've got an IPOD Nano, it's really useful, sometimes I just play different music depending on what I feel like I'm into, sometimes I declare myself to have an 'album of the month' that I listen to, I'll play it every day on that month, until the end of the month, just to get a feel for it, if I haven't listened to it much before."
"So what I did as a test, I [laughs] - there were a hundred exercises of about two lines in this book [of piano exercises], and I wanted to see like how [laughs] how many I could get through in five minutes. [It gave me a] sense of achievement".
"I liked the fast and rhythmical first, I think as a novice I liked to see how fast I could play something, I liked fast music."

1.2 (aesthetic):
"I like harmony more than melody. I can remember a passage from the Pearl Fishers, a duet, and can hear both voices singing at once. The rules of harmony are mathematical".
"I like to recognise the sheer beauty of the music, as opposed to a specific verbal emotion."
"I like some rock music, for example The Who, 'Won't get fooled again', with the synthesiser in that … it's the patterns that I would see as the common factor between Bach and The Who … a very high degree of structure and pattern, and developing, and yet with no feeling of mechanicalness about it."

1.3 (belonging):

“When someone’s released a classic album and you listen to it, I think ‘I’m part of something great now’”.
“There’s something about pop music, it’s the image, and to do with people of my generation, rather than adults inflicting it on you.”
“I hear a lot of music in church, I go on Sunday and Monday and Thursday as a rule … you can relate to it, sometimes if you're not at that church, if the same song comes along sung by a different congregation, you can relate, it makes you feel a part of something…”.

1.4 (enjoyment):
“Some of the music, I can recognise it as sad, but I enjoy it so much, that the sadness kind of doesn’t make me sad, I just enjoy listening to the music”.
“I have a music collection for my own enjoyment - none of my brothers have”.
“At college and since, I have sung in choirs, it gives another dimension to the music [than piano playing], but I have a similar feeling of enjoyment”.

1.5.1.1 (valency change):
“When I started getting more severe medical problems a few years ago, then I was playing happier music to cheer myself up. Not exactly cheer me up, but keep me on a level.”
“I was having a bad period, and I felt really down and out about myself, and sometimes I would listen to “Automatic for the People” and “Out of Time” as well, and when I listened to Automatic, it made me feel better about things, it was like a silver lining, and that was how I got to like REM so much.”

1.5.1.2 (no valency change):
“I don’t feel happiness, though have other intense emotions, and so don't listen to ‘happy’ music.”
“I am often - not depressed exactly, but sombre - this is my default mood. There are just variations from there. So if music were to change my mood it would be to sadness not to happiness. I haven't tried playing happy music to myself to ‘cheer up’. Sombre is how I am, I would not want to listen to happy music in that case. Exhilaration is the closest I get.”

1.5.2.1 (induces arousal):

"As for exhilaration, I love Beethoven's 6th (Pastoral) - the thunderstorm - the calm after the storm - you can feel all the emotions you want in that music - I love classical music. But also Radiohead for exhilaration - their music has so much power. Also like Pink Floyd, Jefferson Airplane."

"When I used to listen a lot more to music, I used to listen to Chuck Berry, that got me excited."

"They [bands] have to have ideas, even if you don't grasp it exactly. Has to be a bit different from everyday life. Something about the ethos and excitement of it."

1.5.2.2 (induces relaxation):

"If you feel tense or whatever, you want something to calm you down, then you listen to Japanese music."

"Sometimes I feel calmness or serenity, like listening to Gregorian chant, that really calms me, makes me feel at peace."

"If I'm like driving into London, … or anywhere urban where there's a lot of traffic and it's start-stop, I listen to something relaxing like REM or Jack Johnson, something a bit smooth so that I don't get so over-anxious and over-heated up, because there are a lot of idiot drivers out there who get you wound up so easily".

1.6 (mood congruent):

"I find that sometimes if you're feeling very sad or something listening to that kind of music can put you in touch with your feelings, it can help you to access your feelings. You can really feel the feelings instead of their just being there, you can really dwell in that state and deal with it."

"I was a bit of a lonely person when I was younger … I had that emptiness in me, and I was going through those growing up stages, so I listened to the radio and I listened to Heart 106.2 which played all the smooth love songs and kind of sad songs, and it reflected how I felt at the time."

1.7 (movement):

"There have been times when I really felt down about stuff … but when a song came out [by REM] then I would be more able to move to it …"

"Like, its hard to describe, you can imagine [music] makes you feel you are in a certain environment or feeling in a certain way, roaming through the countryside, getting away from your immediate surroundings. But not visually, its more a sort of feeling."

1.8 (performance):
"I'm a lot of the time a restless character, and I got a bit carried away with Queen, because I wanted to perform, to perform on stage and have people adoring me, I just wanted attention, I wanted people to say 'oh look at me', I'm a showman".
"I loved singing [as a child] – I was in the church choir – I learned to dance, and loved anything to do with music and movement."
Note that in addition, one participant stated that he regularly gave amateur piano recitals (of which we have independent confirmation), one regularly plays piano and organ at his local church, and one (a self-taught pianist) plays an hour of "easy listening" background music in the auditorium of a local community centre, four days a week.

1.9 (therapeutic):
"When I have been feeling depressed, I have listened to certain music, and I would claim the music healed me".
"Nothing worked for me - therapy didn't work - but listening to this chap [singer Carl McCoy] actually worked".
"With autism you tend to feel dead a lot of the time - music is the key that unlocks the emotions".

2.1 (appropriate association):
"Gregorian chant has this effect because it is old music - I can imagine being in a monastery - I love old buildings - I feel their age and history and have great respect for them. They give a feeling of sanctuary".
"You can either listen to the music, you can listen to the words, or you can listen to both, and you can think a song is nice, because even if the lyrics are telling a sad story, the music can make you think of something you like, and in this case trees and countryside and fresh air".
"I was thinking about other things I remember being particularly moved by - one piece of music that comes to mind is Beethoven's Ninth. I remember listening to a recording of that, and at the same time reading the notes of how Beethoven was conducting the first performance of that,

but couldn't hear the audience - gosh, I feel choked up now, just thinking about it - he couldn't hear the audience's applause behind him …”.

2.2 (emotional expressiveness):
“I used to listen to Beethoven, the Eroica, and that’s very sad, the funeral march”.
“Groups like the Beatles sound quite happy, though ‘All you need is love’ is a dreary tune, like a dirge, though it’s supposed to be about love”.
“Therapy didn’t work - but listening to this chap [a singer], actually worked. It was because he was even darker and angrier than I was, and it just seemed to get underneath and cut it [the anger] off at the roots”.

2.3.1 (relaxing or slow):
“I do like slow movements as well, that are less rhythmical, in which the beat is much slower. Like you have a slow movement of a sonata. It took me more of a while to acquire the taste for slow movements”.
“If I’m trying to relax I’m not going to listen to something that’s loud or fast paced or … I’m going to chill out or something”.
“Yes it can be quite relaxing listening to slow things, late at night I mean, I suppose if I want to read my emails late at night as I often do, it’s probably better to play that than something quite lively, which will probably have me getting up off the chair and pacing round the room”.

2.3.2 (power, speed, beat):
“[I like] the brass, they add a lot of power to music…”.
“The music from the ‘80’s [that he liked] had a lot of beat and rhythm to it. I don’t like slow, quiet music at all, it tends to be a bit boring. I sometimes like listening to the Spanish music, or I like German music as well, because that’s more go-ey … that has a lot of beat in it”.
“The classics are very powerful, Tchaikovsky, these are great composers aren't they, they really are great”.
“With punk there’s a kind of raw energy [reason for liking it]”.

2.4.1 (familiarity a requirement):
“With music you get to know a piece of music, what's ahead, if you like it, you enjoy it, but conversations are unpredictable [reason to prefer music to conversation]”.

[Reason given for liking certain music]: “probably the sameness of it. I like the Rolling Stones music. if I listen to music I'm not familiar with I’m not very happy with that. You have to keep listening and listening to music. I always used to listen to ‘The Sound of the Sixties’ on Saturdays, but my routine’s changed now, I do my shopping then. Routines seem to have taken over a lot of things in my life”.
“[I like] familiarity - hence I like Classic FM, rather than Radio 3 which tends to be unfamiliar”.

2.4.2 (variety a requirement):
“I need a change sometimes, need to listen to something a bit faster and happier. Change of tone. And I don’t want to get obsessed!”
“But he has his own unique style, he doesn’t really have a particular sound, which is a good thing, because if you sounded the same all the way through people would lose interest. You’ve got to be versatile”.
“I’ve got an IPOD Nano, it’s really useful, sometimes I just play different music depending on what I feel like I’m into … it’s got about 850 songs on it at the moment, some of them I don’t listen to because I’m a bit let down by some of them, I listen to mainly rock and indie and sometimes a bit of pop as well”.

2.5 (melody and harmony):
“I like harmony more than melody. I can remember a passage from ‘the Pearl Fishers’, a duet, and can hear both voices singing at once. The rules of harmony are mathematical”.
[Speaking of what he likes about music] “the grandeur, that kind of thing, and the melodic qualities, and the fact that it has quite adventurous harmonies … some of the songs, you have to really get caught in the melodic qualities of it, the dramatic effect gets you quite excited”.
“And again with Abba and the Beatles, there was good musicianship, and good melodies”.

2.6 (message):
“At times - it’s the bands that are a bit different - and what they have to say, their philosophies. They have to have ideas, even if you don’t grasp it exactly. Has to be a bit different from everyday life. Something about the ethos and excitement of it”.

“I liked Bob Dylan. Lyrics were his strong point. He hasn’t got a good voice, but he puts it over with such conviction, the conviction is really important, as well as the lyrics, and with Carl McCoy of the Nephilim, that’s what’s important, and the belief that he’s 100% in what he's singing, and that comes across in the voice”.
“I was feeling a bit gloomy, not depressed, just sad that I was being bullied a lot, and I couldn’t do anything much about it, and the mood from REM helped me to … it was as if there was a message from them, saying ‘don’t worry about things, we’ll play this music and I hope it makes you feel better’”.

2.7 (social connections):
“I’m reminded when I hear music, of people who like it, I've a friend who's a good New Order fan and he introduced me to New Order, when I listen to them it can remind me of him. My sister likes particular music, when I listen to that I'm reminded of her”,
“What started me off collecting records was we had a friend, and he introduced me to Beano’s in Croydon which is a second-hand record shop, and it started off from there”.
“I went to Australia when I was 20, and my friend in a record shop saved me a box of singles, when I heard some new groups”.
“Me and my mates Tony and Simon, we are very much into Phil Specter”.

2.8 (structure and pattern):
“I definitely prefer music [to conversation]. It depends on who is speaking - certain voices grate - loudness, lack of any set pattern - it’s just noise. With music you get to know a piece of music, what’s ahead, if you like it, you enjoy it, but conversations are unpredictable. There is no pattern to them. It’s like playing the piano out of key: the wrong notes at the right time”.
“It [music he liked] reflected the way I was feeling, and it wasn’t just one song, it was lots of songs, and I thought it’s not just the way these songs are, it’s the order in which they’ve been put on the CD, if it was in a different order it wouldn’t sound the same, it wouldn’t have the same feel”.

2.9 (technical competence):
“They do know how to play their instruments, and I do like musicians who know how to play”.
“With Culture Club, Boy George, you can understand what he’s singing, George Michael, you can understand what he’s singing about, but a lot of the heavy rock, like AC/DC, Motorhead,

and Thin Lizzie or whatever, they just shout, and you can hardly hear with all of the drumming and the guitars as well".

"I like The Stranglers a bit more than The Clash, because The Stranglers had some good keyboard … and again with Abba and the Beatles, there was good musicianship".

2.10 (timbre):

"He [McCoy] has a very deep resonant voice".

"The legendary Enrico Caruso, he had such a rich voice".

"I mean the fact that it was quite a powerful organ, with, like, 32-foot reed stops which has a very deep sound, almost an octave below the piano, that pitch is almost outside human perception, but the power of it, you know, really adds resonance, it sort of gets right into my roots, if I was a tree it would get right into my roots". [Speaking about some organ music that he liked].

APPENDIX 4.1: Musical items from Quintin et al.

The identity of pieces from which they were chosen, and the associated emotions, were as follows:

1. “Migration”, by Carlos Nakai and Peter Kater. (Sad).

2. “Anthem” by Susanne Ciani. (Happy).

3. Extract from movie score for “Spellbound”, by Miklos Rozsa. (Scary).

4. “Phantazia” by Noel Pointer. (Sad).

5. Extract from movie score for “Marnie”, by Bernard Herrmann. (Scary).

6. Extract from Mahler’s Symphony No 5, Adagietto. (Sad).

7. Extract from movie score for “Vertigo”, by Bernard Herrmann. (Scary).

8. “Variations of the Wayfaring Stranger” by Lucas Foss. (Happy).

9. “Rainstorm” by Noel Pointer. (Happy).

10. Faure's Quintet No 1 in D minor (molto moderato). (Sad).

11. “Rite of Spring” by Stravinsky. (Scary).

12. Prelude no. 15, by Bach. (Happy).

APPENDIX 4.2: Words or phrases used by the autism and control groups

Words used by the autism group:

Excited, tranquil, chilled out, tense, feel a buzz, exhilarated, relaxed, serene, peaceful, soothed, invigorated, anxious, calm, lively.

Words used by the pilot control group:

Bright, magical, happy, thrilled, sad, dramatic, pleasure, escaping, thoughtful, cheerful, longing, scared, wistful, hopeful.

These were presented in randomized order as a check sheet for participants in the word response study reported in chapter 5, as follows ("item number" refers to the musical items, in the pseudo-randomized order in which they were presented):

Item no:	1	2	3	4	5	6
excited						
tranquil						
bright						
chilled out						
magical						
tense						
feel a buzz						
exhilarated						
happy						
thrilled						
relaxed						
serene						
sad						
peaceful						
dramatic						
pleasure						
soothed						
invigorated						
anxious						
escaping						
thoughtful						
cheerful						
longing						
scared						
wistful						
calm						
hopeful						
lively						

APPENDIX 4.3: Lists of word "bundles"

Bundles 1 through 6 correspond in order to tracks 9, 1, 5, 11, 8 and 10 in appendix A.

BUNDLE ONE	BUNDLE TWO	BUNDLE THREE
happy movement irritated joyful fun sunshine	relaxed peaceful melancholy calm sad	alert anticipation suspense, tension eager excitement sense of danger
BUNDLE FOUR	BUNDLE FIVE	BUNDLE SIX
threatened oppressed Scared confusion, bewildered slipping and falling	lively dancing energised upbeat adventurous exuberant	grief and despair intensity calmness, tranquility reflective, thoughtful loss

APPENDIX 4.4: Instructions to participants

I am going to play you six items of music, and then ask you about your responses to them. To begin with, I would like you to take a few minutes to look at this list of words. If you come across a word which you don't understand, please ask me when you come to it and I will try and explain it.

Please now read the words to yourself, not aloud, and tell me when you have finished.

I will now play the music, one piece at a time. While I am playing the first piece, please look again at the words and decide if any of them describes the way the music makes you feel, or describes a thought that comes to you when listening, even if momentarily or fleetingly. If there are any words like this, please put a tick next to them in column 1, which is this one. You can do this either during the music or after it has finished. It is quite all right if you decide that none of the words applies, and in that case just leave the column blank and tell me when you are ready to move on to the next item of music.

Please note that I only want you to record the thoughts or feelings that the music evokes in you. I am not looking for a description of the music itself, but of what is going on in your mind in response to the music, using the words provided in the list. If you have a thought or feeling that is not on the list, please tell me what it is.

Now I will play you the next item, and I would like you to do the same in column 2. Please remember that the feeling should be one that **you have yourself** – it may be that this is not the feeling you see expressed by the music. If for example you think the music sounds sad, but you yourself feel pleasure and not sadness (perhaps because it has pleasant memories for you), you should mark 'pleasure' and not 'sad'.

Now please look at the following set of six "bundles" of words that might be used to describe music. I am going to play you six more items of music.

This experiment is different from the last one. The bundles represent typical words used by an earlier group of people to describe the way they felt about the different items of music. I would like you to use your judgement to decide which bundle you think was the one used by the group to describe that piece of music. The words in the bundle don't necessarily have to represent the way the music makes you feel, or a thought that the music evokes in you.

You do not have to choose different bundles for each item of music. It is fine if you want to choose the same one to describe two or more items.

APPENDIX 5.1: Calculation of the effects of sound stimuli

The program devised by Davis which took the readings of GSR resistance in kilohms, sampled the values and stored them on computer at the rate of two readings per second. The program also recorded the exact time at which each GSR reading was taken, and also whether there was a sound stimulus being presented at that point (and if so it identified the code number of the stimulus) or whether there was silence at that point.

The data were analyzed on a participant by participant basis, ie the effect sizes were calculated for each individual and yielded two quantities: the GSR effect of the music extracts (based on the combined effects of all twelve of them), and the GSR effect of the environmental noise extracts (similarly, calculated on the basis of the combined effect of all six noise samples).

In order to analyze the data, I wrote a simple Windows Excel program that automatically analysed the data for each individual. The program was applied separately to the data recording response to music, and to noise stimuli.

The first step executed by the program was to locate each point at which there was a change of stimulus, ie where either the sound (music or environmental noise) ceased and a period of silence began, or vice versa (I call this a “transition point”). For each transition point, the program also measured the GSR readings taken at two other points, located at five seconds before, and five seconds after the transition point (ie at ten sample readings before and after the transition point). I call these signal readings the “bracket points” for that transition point. There was no particular reason for choosing an interval of five seconds, but the choice seemed reasonable as giving a sampling of the reaction to the sound or silence stimulus that would provide a measure of the short term trend of the readings that would be both large enough to minimise the distorting effects of random noise, and short enough to represent the trend prevailing over a short time period, a trend which might fluctuate substantially when listening to music whose characteristics might vary over a substantially longer interval.

For each transition point, the program averaged the readings at the two corresponding bracket points, and subtracted them from the reading at the transition point. For technical reasons, in order to give a result in kilohms per second, this result was divided by a scaling factor of 2.5 (the reason for this is explained below). The program then calculated the absolute value of the resulting number for each transition point, and finally, took the overall median of these absolute values. The median was adopted, since it this statistic is known to be less susceptible to the influence of possible artefacts or other causes of extreme values, than the use of a simple average.

Algebraically, if the reading at the transition point at time t was R_t, where t is measured in seconds, then readings were taken at the bracket points corresponding to R_{t-5} and R_{t+5} and the quantity $|R_t - (R_{t-5}+R_{t+5})/2|$ was calculated for each value of t corresponding to a transition point, and then the effect size was finally calculated as
effect size = median$\{|(R_t - (R_{t-5}+R_{t+5})/2)|/2.5$, taken over all t at transition points$\}$. The units of GSR measurement were in kilohms, so the units of this effect size measure were in kilohms per second.

The rationale for this is that in the absence of any effect of sound vs. silence, one would expect the slope of the trace line at the transition point to be continuous before, during and after the transition point. The simplest and most parsimonious hypothesis is that the change from music to silence or vice versa has a first order linear effect on the slope, with a zero effect size corresponding to a zero change of slope. It is easy to see that if R_t were a linear function of time t (i.e. if the transition had no effect), then
$R_t - (R_{t-5}+R_{t+5})/2$ would be zero. In fact, if the trace of resistance against time is visualised as a graph with time as the x-axis and resistance as the y-axis, then
$R_t - (R_{t-5}+R_{t+5})/2$ is the vertical distance between the point represented by coordinates
(t, R_t), and the midpoint of the line joining the points $(t-5, R_{t-5})$ and $(t+5, R_{t+5})$.

The estimate of rate of change of R with time during the period between the first bracket point and the transition point is clearly, by definition of rate of change, equal to
$(R_t - R_{t-5})/5$, in units of kilohms per second. The rate of change of R in the second time period is $(R_{t+5} - R_t)/5$, in units of kilohms per second. The difference between these two, which by a simple calculation is $(R_t - (R_{t-5}+R_{t+5})/2)/2.5$, is the difference in the slopes of the traces immediately before and after the transition point, and this was taken as the definition of the effect size of the sound/silence or silence/sound transition at that point.

APPENDIX 5.2: Key for words coded numerically in Figure 9.

1	excited
2	tranquil
3	bright
4	chilled out
5	magical
6	tense
7	feel a buzz
8	exhilarated
9	happy
10	thrilled
11	relaxed
12	serene
13	sad
14	peaceful
15	dramatic
16	pleasure
17	soothed
18	invigorated
19	anxious
20	escaping
21	thoughtful
22	cheerful
23	longing
24	scared
25	wistful
26	calm
27	hopeful
28	lively