CHAPTER 2
MAPPING EMPIRICAL RESEARH STUDIES ON MUSIC AND EMOTION

2.1 INTRODUCTION

This chapter focuses on empirical research conducted in relation to music and emotion. It presents existing methods as well as potential methods that could be used in future research. It furthermore contains a description of the studies and measuring instruments upon which the research design of the study described in this research report was based, as well as a discussion of other method-related decisions that were made while the research method was designed. This chapter therefore answers the first two sub-questions of this study.

1) Which existing questionnaires and tests can be utilised to establish profiles of individual listeners and interpret data concerning the emotional content of listening experiences?
2) What kinds of continuous testing are used when musical listening experiences are explored, and what kinds of continuous testing are used when emotions are explored?

The chapter excludes any discussions on philosophies or theories in research on music and emotion, and focuses only on methodologies, since the current study is a methodological study. The chapter is concluded with a summary.

2.2 MAPPING OF EXISTING AND POTENTIAL RESEARCH

Figures 2.1 – 2.9 form a representation of what is termed in this report 'the methodological map'. The map represents an attempt to offer an overview of methods used in studies of which the focus was music and emotion, or in some cases, emotional experiences to other stimuli (like facial expressions or colours). The purpose is to show the ways in which other researchers measured emotional responses. This map was compiled by studying research reports published in the field.

The map was then augmented by considering other potential methods which have not yet been pursued as far as the current researcher is aware. These potential methods are presented in italics in the map. The researcher offers no claim as to the completeness of the map, and is aware that this map can also be drawn in many other ways. Examples of existing studies are provided in Table 2.2 to support the overview. The numbers provided in the map are also represented in the tables.

Figures 2.2 - 2.4 represent methods for compiling profiles of participants. The map was divided over three figures since all of the information could not fit into one figure. Figure 2.1 is a summary in terms of the kinds of profiling methods, and shows how the three parts in Figures 2.2 to 2.4 are related.

Figure 2.1 Profiling methods: summary

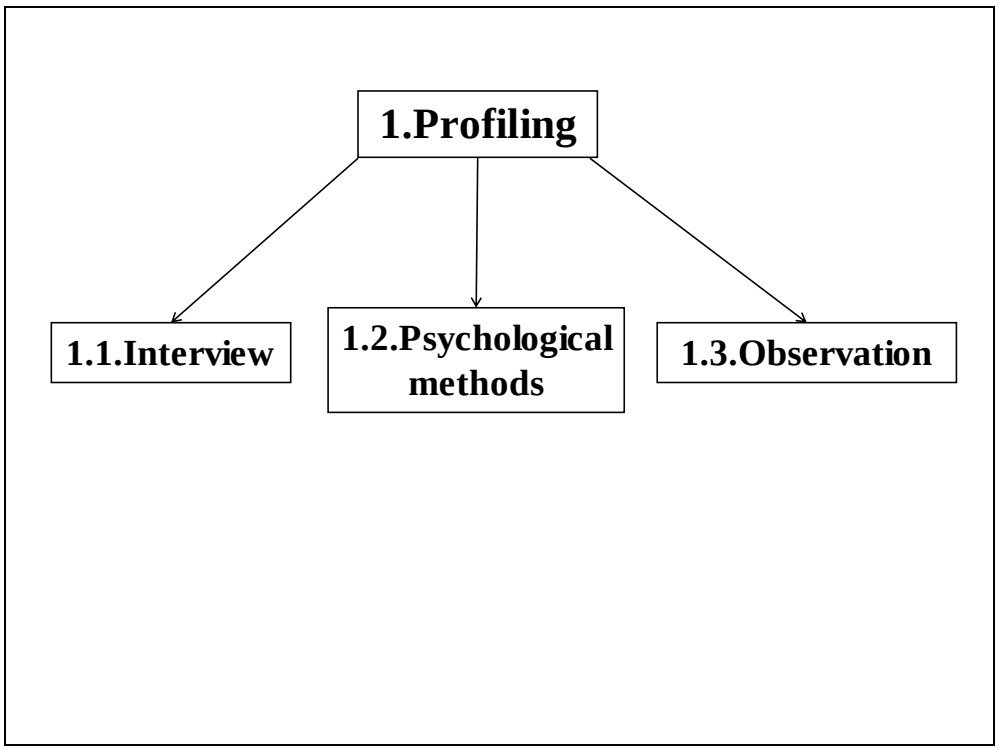


Figure 2.2 Profiling methods 1/3

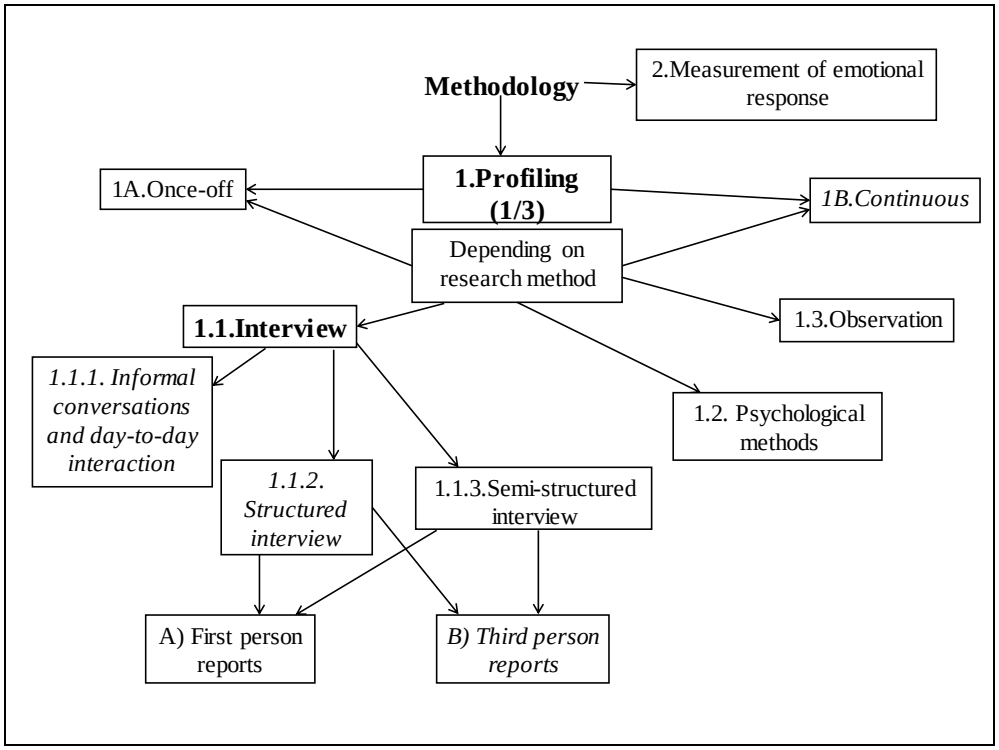


Figure 2.3 Profiling methods 2/3

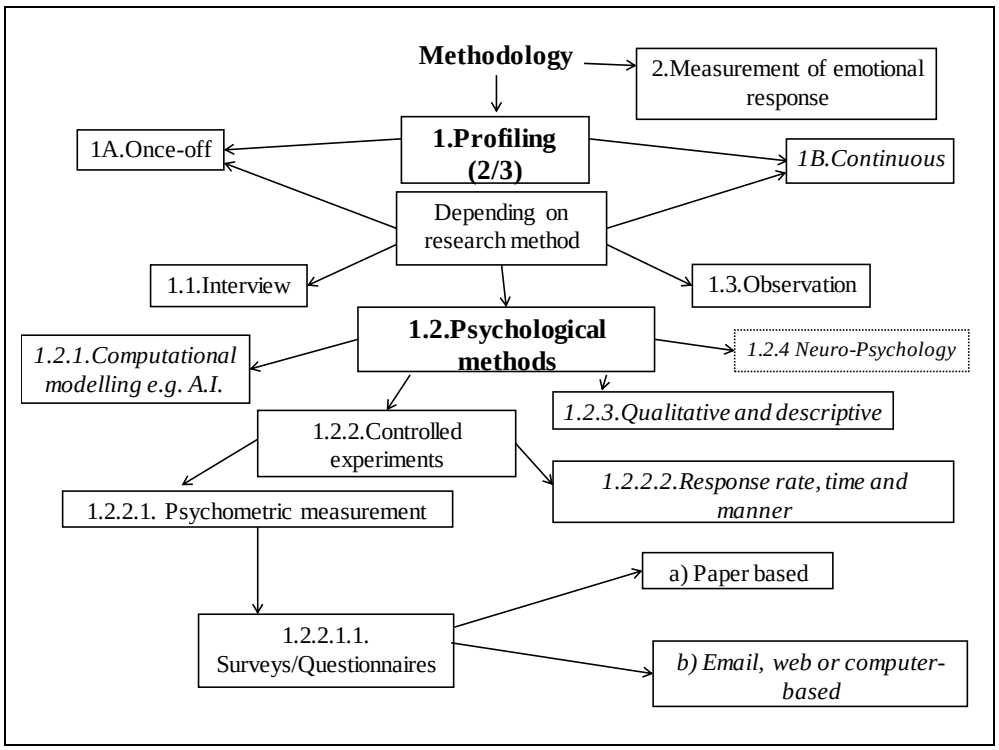
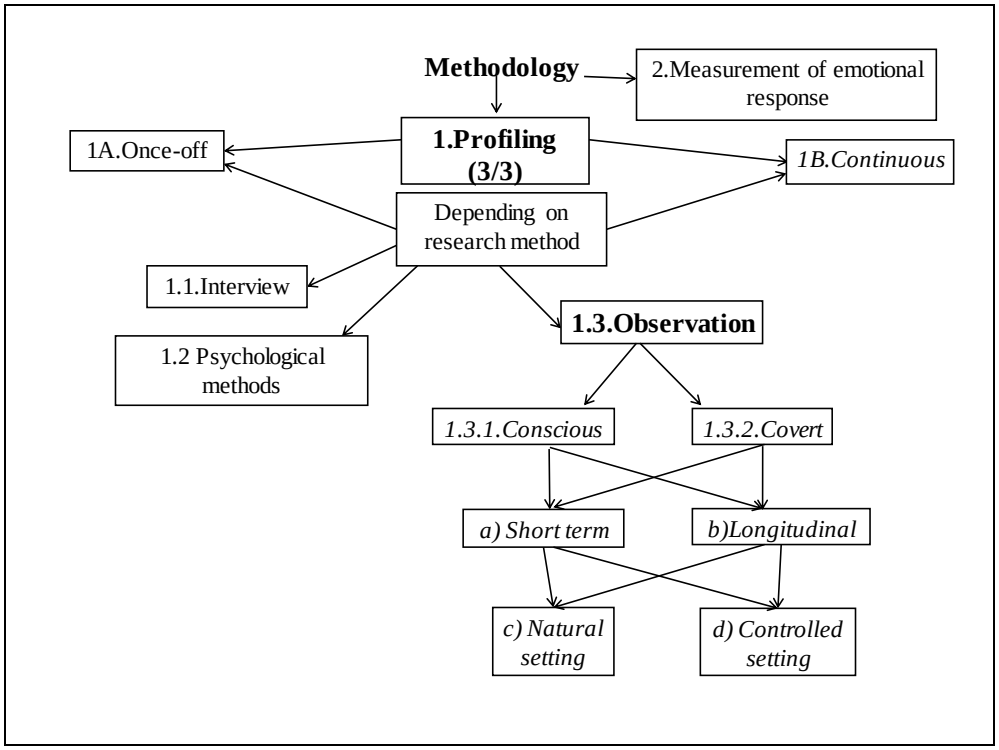


Figure 2.4 Profiling methods 3/3



Figures 2.6 – 2.8 represent methods to measure emotional response. The map was also divided over three figures. Figure 2.5 is a summary in terms of the kinds of methods of measuring emotional response.

Figure 2.5 Methods to measure emotional response: summary

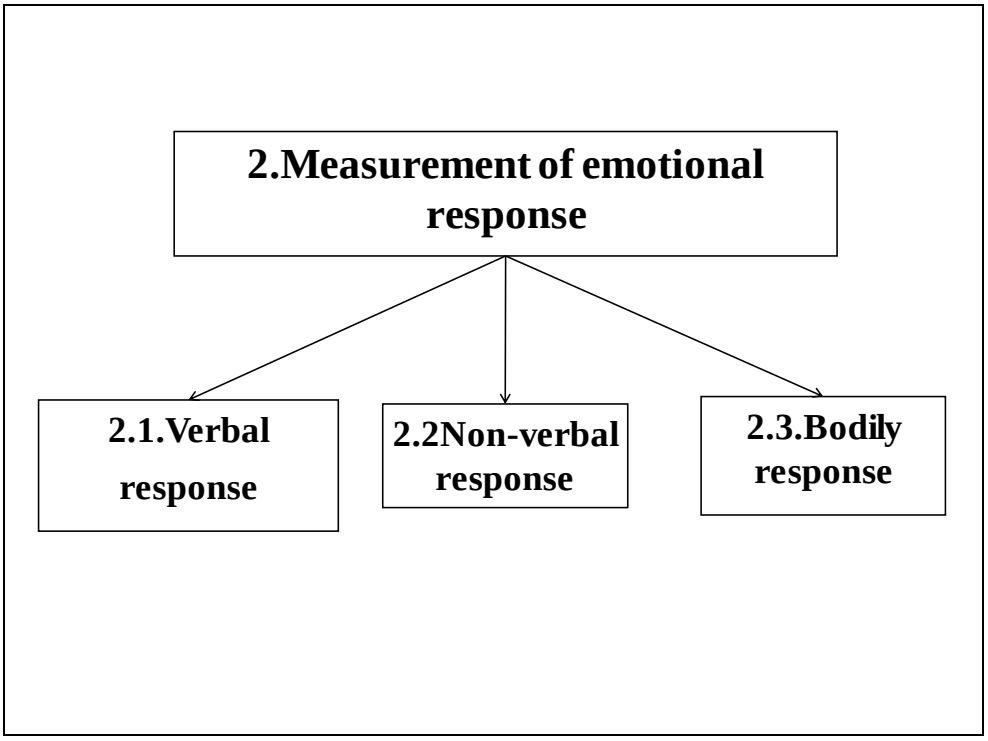


Figure 2.6 Methods to measure emotional response 1/3

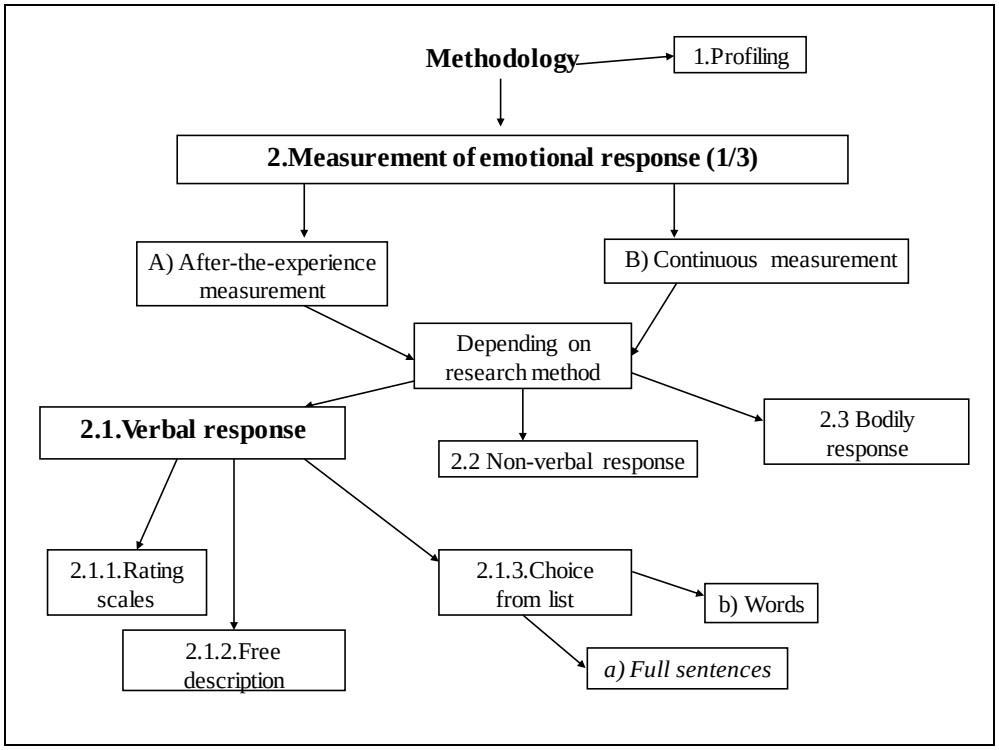


Figure 2.7 Methods to measure emotional response 2/3

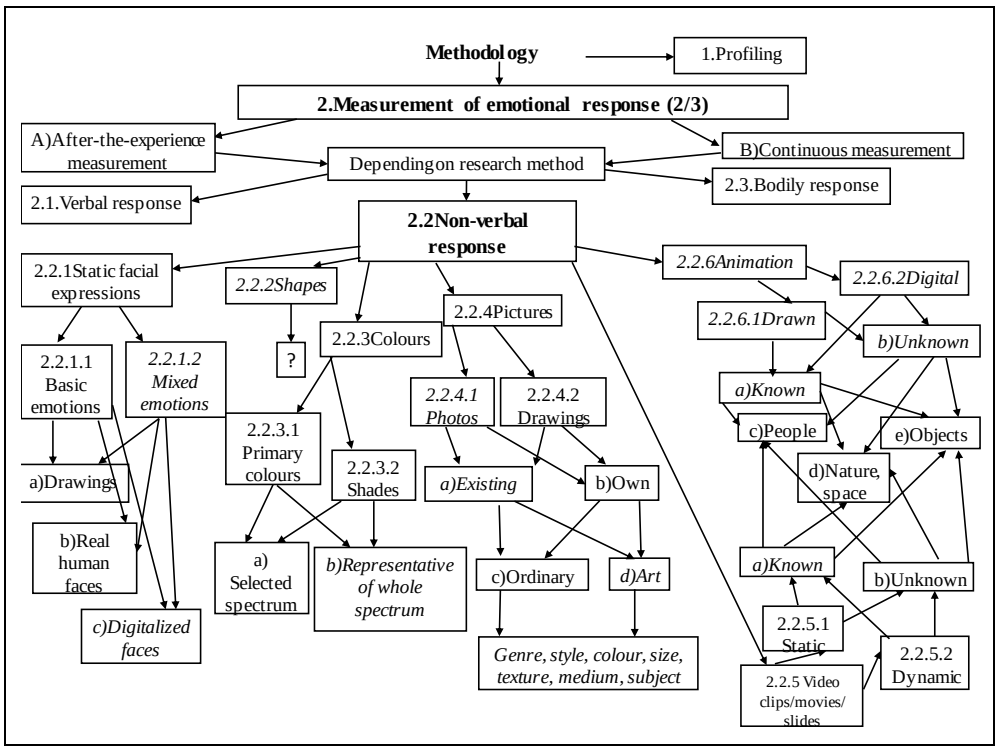


Figure 2.8 Methods to measure emotional response 3/3

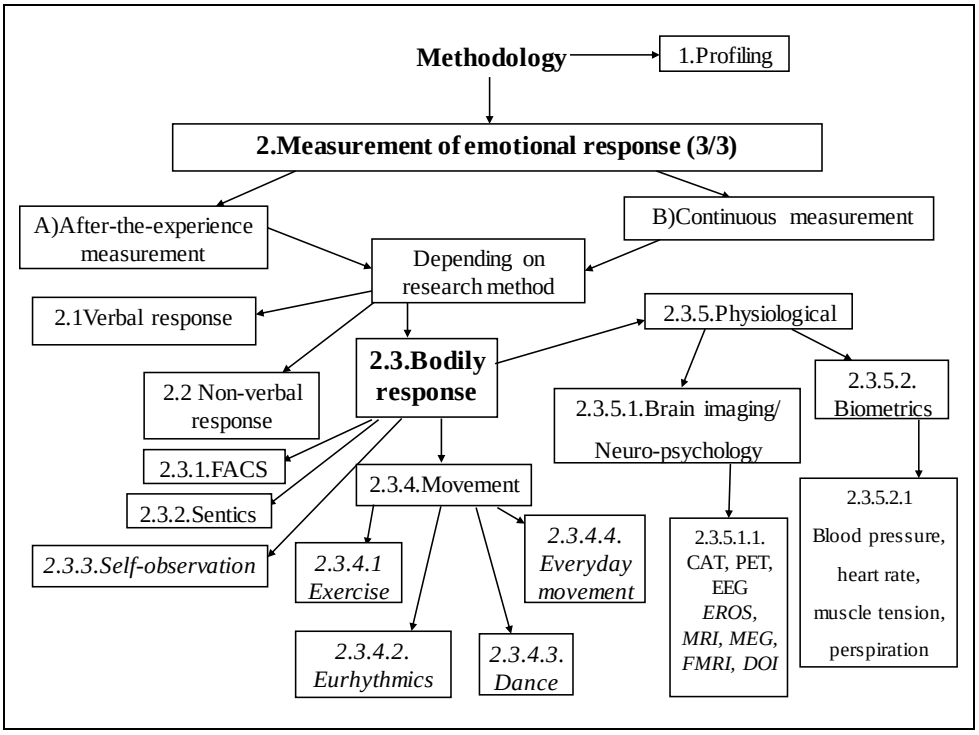


Table 2.1 offers descriptions of the acronyms in 2.3.1 and 2.3.5.1.1 of Figure 2.8.

Table 2.1 Acronyms

FACS	Facial Action Coding System which refers to a procedure to analyse human facial expressions.
CAT	Computed Axial Tomography, which is used for viewing and measuring brain injuries.
DOI	Diffuse Optical Imaging, which measures optical absorption of a metallo-protein containing oxygen and iron.
EROS	Event Related Optical Signal, which measures changes within the active areas of the brain.
PET	Positron Emission Tomography, which measures emissions from radioactively labelled metabolically active chemicals that have been injected into the bloodstream.
MRI	Magnetic Resonance Imaging, which creates a complete image of the brain's anatomical detail and structure.
MEG	Magneto encephalography, which measures neurological feedback and functions of the different parts of the brain.
FMRI	Functional MRI, which does the same as MRI, but also measures changes in blood flow related to neural activity.

Figure 2.9 represents a summary of musical stimuli that can be used in studies.

Figure 2.9 Musical stimuli: summary

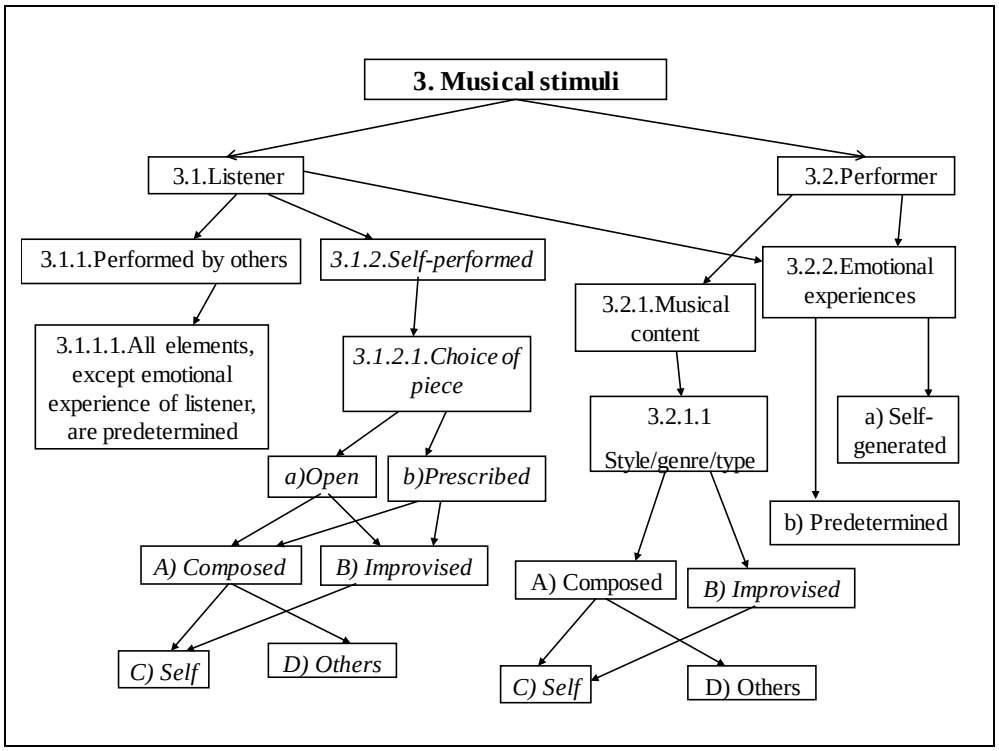


Table 2.2 contains examples of existing studies that employ some of the methods that have been indicated in the methodological map. The method used and reference of the study is provided along with a short description of the study. The number next to the description of the method is the number used in the methodological map.

Table 2.2 Existing studies

Method	Reference	Description
1. Profiling 1.1 Interview 1.1.3. Semi-structured interview A) First person reports	Green, 2002	Green conducted semi-structured first person interviews with popular musicians in order to determine how they learned.
	Thompson, 2004	The study was conducted to determine aural skills and behaviour among undergraduate music

		students. She conducted four aural tests and semi-structured interviews with each student over a four-week period (continuous profiling).
1.2 Psychological methods 1.2.2. Controlled experiments 1.2.2.1.Psychometric measurement 1.2.2.1.1. Surveys/ Questionnaires a) Paper based	Balkwill *et al.*, 2004	A demographic questionnaire was used in the study.
	Ryan and Andrews, 2009	A questionnaire was given to semi-professional choral performers to determine the cause, frequency and severity of performance anxiety.
2. Measurement of emotional response 2.1. Verbal response 2.1.1. Rating scales	Balkwill and Thompson, 1999	Western listeners were asked to rate the degree of four different emotions in Hindustani raga samples.
	Balkwill *et al.*, 2004	Japanese listeners were asked to rate the expression of three emotions in music from three different tonal systems.
	Tan *et al.*, 2006	Rating scales were used to compare the experience of complete vs. patchwork (fractured) compositions for both trained and untrained musicians. The scales included questions about the emotional content of the music.
	Patrick and Lavoro, 1997	Students viewed sixty colour photographic slides and rated their emotional responses to each slide using the positive and negative affect schedule.

	Davis *et al.*, 1995	Participants were shown a series of slides and then used a rating scale to provide an emotional response to each image.
2. Measurement of emotional response 2.1. Verbal response 2.1.2. Free description	Aiello *et al.*, 1990	Trained musicians were asked to describe a piece of music by using prescribed musical terminology, but with the freedom to add their own interpretation using words not included in the prescribed list.
	Tan and Kelly, 2004	Groups of formally and informally trained musicians were asked to listen to various pieces of music, and then visually describe their experience of the music using a self-made drawing. They then had to provide a written explanation of their drawing.
	Sloboda and Juslin, 2001:74, 84	The authors list free description as a common way to measure emotional responses, and also reference a study conducted by Pike. In Pike's study, musically untrained participants were instructed to listen to different pieces of music, and then write down their emotional responses.
	Gabrielsson and Lindström, 2001:225	The authors list free description as a means to report emotional experience and mentioned that it was used in studies conducted by Gilman and Downey, respectively.
2. Measurement of emotional response 2.1. Verbal response 2.1.3. Choice from list b) Words	Bresin and Friberg, 2000	Participants were asked to select a word from a list in order to identify the emotional expression of the music sample used in the study.
	Farnsworth, 1953	Participants listened to musical phrases and provided an emotional response using the Hevner

		word list.
	Sloboda and Juslin, 2001:74	The authors list word checklists as a common way to measure emotional responses.
	Gabrielsson and Lindström, 2001:225	The authors also list choice among terms as a means to report emotional experience, and referenced a study conducted by Gundlach which used this method.
	Gabrielsson and Lindström, 2001:230	In studies conducted by Hevner, participants were required to mark as many terms as they found appropriate for a musical piece.
2.2 Non-verbal response 2.2.1 Static facial expressions 2.2.1.1 Basic emotions a) Drawings	Da Pos and Green-Armytage, 2007	Their study investigated the relationship between facial expressions, colour and basic emotions. They used drawings of facial expressions displaying six basic emotions which the participants had to match with colours representing their best judgment of which facial expression matched with either one or three colours.
	Zentner, 2001	Students were asked to match three facial expressions to a list of colours, as well as to indicate their colour preferences. The facial expressions used in the study were drawings of expressions of happiness, sadness and anger.
2.2 Non-verbal response 2.2.1 Static facial expressions 2.2.1.1 Basic emotions b) Real human faces –	Ekman, 1978:100	Ekman described a study by Ekman, Sorenson and Friesen where they showed photographs of the six different emotional facial expressions to people in the USA, Japan, Chile, Argentina, and Brazil. Observers had to select one of six emotion words for each photograph they saw.

photographs		
2.2 Non-verbal response 2.2.3 Colours 2.2.3.1 Primary colours a) Selected spectrum	Da Pos and Green-Armytage, 2007	They investigated the relationship between facial expressions of basic emotions and colour, using selected colours.
	Kaya and Epps, 2004	Colleges students were asked to indicate their emotional responses to certain colours.
	Hemphill, 1996	This study investigated students' choice of favourite colour, the colour they wore most, and the emotional responses they related to the colours used in the study.
	Zentner, 2001	Students were asked to match three facial expressions to a list of colours, as well as indicate their colour preferences.
	Hoshino, 1996	Participants listened to samples from both Western and Japanese music and provided an emotional response using the sixteen colours of the Matsuoka colour table.
2.2 Non-verbal response 2.2.3 Colours 2.2.3.2 Shades a) Selected spectrum	Da Pos and Green-Armytage, 2007	They investigated the relationship between facial expressions of basic emotions and colour, using selected shades as well.
	Kaya and Epps, 2004	College students were asked to indicate emotional responses to certain shades of primary colours.
	Zentner,	Students were asked to match three facial expressions to a list of colours, as well as indicate

	2001	their colour preferences. The list also included certain shades.
	Hoshino, 1996	Participants listened to samples from both Western and Japanese music and provided an emotional response using the sixteen colours of the Matsuoka colour table, which included shades.
2.2 Non-verbal response 2.2.4 Pictures	Ekman, 1993	He mentions the use of slides as a test of emotional response. The kind of pictures used in the study was not specified.
2.2 Non-verbal response 2.2.4 Pictures 2.2.4.2 Drawings b) Own c) Ordinary	Tan and Kelly, 2004	Groups of formally and informally trained musicians were asked to listen to various pieces of music, and then visually describe their experience of the music using a self-made drawing.
	Oh, 2006	Children were asked to draw their emotional responses to musical stimuli. The stimuli were short extracts from classical music pieces that were supposed to correspond to the emotions of happiness, sadness, anger and fear.
	Gabrielsson and Lindström, 2001:228	The authors mention a study conducted by Smith and Adams where children listened to six ascending intervals and were required to represent them in drawings.
2.2 Non-verbal response 2.2.5 Video clips/movies/slides	Ekman, 1993	He mentions the use of video clips as a test of emotional response.
2.2 Non-verbal response	Patrick and Lavoro, 1997	Students viewed sixty colour photographic slides and rated their emotional responses to each slide using the positive and negative affect schedule.

2.2.5 Video clips/movies/slides 2.2.5.1 Static b) Unknown; c) People; d) Nature, space; e) Objects		The contents of the slides were specified in the study. The slides were colour photographic slides, depicting a variety of people (men, women and babies), landscapes, animals (wild or domestic), and objects (a plane, or household objects).
	Davis *et al.*, 1995	Participants were shown a series of slides of various people, objects, animals, etc. and then used a rating scale to provide an emotional response to each image.
2.2 Non-verbal response 2.2.5 Video clips/movies/slides 2.2.5.2 Dynamic b) Unknown	Ekman, 1978:100	In an experiment, stress-inducing films were shown to students in the USA and Japan. Part of the time the students watched the film on their own, and part of the time the person watched while talking about the experience with a research assistant from the person's own culture. Measurements of their actual facial muscle movements were captured on video. The content of the video the students watched, were not specified in the reference.
2. Measurement of emotional response 2.3 Bodily response 2.3.1 FACS	Vanger *et al.*, 1998	Photos taken of participants were digitally blended to create an 'average' face for expressing emotions. FACS encoding was used to create the facial expressions and also to evaluate the digital faces afterwards.
2.3 Bodily response 2.3.2 Sentics	Sloboda and Juslin, 2001:91	The authors mention a study conducted by Waterman in which participants listened to pre-recorded pieces of music. They were asked to press a button each time they experienced an emotional reaction or 'felt something'. They were also required to explain their reactions afterwards.
	Gabrielsson	The authors mention a study conducted by

	and Lindström, 2001:226	Nielsen, where participants had to press a pair of tongs in proportion to the perceived tension in the music. The pressures were registered on a polygraph showing pressure curves.
	Gabrielsson and Lindström, 2001:227	A study was conducted by Clynes to investigate tension using non-verbal response. Clynes introduced finger pressure on a sentograph in order to determine how participants expressed different emotions.
2.3 Bodily response 2.3.4 Movement	Becker, 2001:145-146	The author provides observations of Sufi Muslim musicians that can be described as swaying, rising and turning in place as they perform, coupled with intense feelings of ecstasy.
	Becker, 2001:149-150	The author describes observations of Pentecostal worship services that often involve movement such as hand-waving, hand-clapping and foot-stomping.
	Scherer and Zentner, 2001:377	The authors referenced a study by Todd suggesting that expressive sounds induce a perception of self-motion in the listener.
	Sloboda and Juslin, 2001:88	The authors mention a study where young children break out into motion when they hear certain types of music.
2. Measurement of emotional response 2.3 Bodily response 2.3.5 Physiological 2.3.5.1 Brain imaging / Neuro-psychology	Peretz, 2001:108-109	Peretz refers to a study by Sergent and colleagues where PET scans where used together with visual representations of faces to identify the areas of the brain responsible for processing facial identity and facial emotion.

2.3.5.1.1. CAT, PET, EEG		
	Peretz, 2001:116	Peretz refers to a study by Peretz and Gagnon in which a subject, who suffered brain damage and had trouble identifying music, was asked to judge melodies as happy or sad, and familiar or unfamiliar. CAT scans of the brain were used to determine if different areas are responsible for identification and emotional classification.
	Peretz, 2001:119	Peretz mentions several studies where EEG measurements of the brain were taken using electrodes attached to the scalp while the participant is listening to music. The resultant output was shown to be related to the structure of the music.
2.3 Bodily response 2.3.5 Physiological 2.3.5.2. Biometrics 2.3.5.2.1 Blood pressure, heart rate, muscle tension, perspiration	Sloboda and Juslin, 2001:74	The authors mention the work done by William James, who claimed that emotion was basically a perception of internal bodily changes. Researchers have used physiological indices to measure emotion, such as blood pressure, heart rate, muscle tension, and perspiration.
	Scherer and Zentner, 2001:374-375	The authors refer to a few studies. A study by Ornstein and Sobel suggested that music evokes a calmness that decreases blood pressure, reduces stress, and affects the immune system. An overview by Bartlett reviewed a number of studies that measured the physiological effects of music, including pulse rate, muscle tension, and skin temperature. A study performed by Krumhansel describes a method according to which students listened to pieces of music which were expected to

		induce specific emotional states, while their physiological parameters such as heart rate, blood pressure, skin conductance and temperature were measured.
3. Musical stimuli 3.1 Listener 3.1.1 Performed by others	Balkwill and Thompson, 1999	Western listeners listened to recordings of Hindustani raga samples.
	Balkwill *et al.*, 2004	Japanese listeners listened to recordings of Western, Hindustani and Japanese music.
	Bresin and Friberg, 2000	Participants were asked to identify the emotional expression of musical samples that were computer-generated and manipulated.
3.2 Performer 3.2.1 Musical content 3.2.1.1 Style/genre/ type a) Composed; d) Others	Gabrielsson and Juslin, 1996	Performers were instructed to individually play different monophonic pieces, each representing different musical styles with varying emotional characters. The pieces also had to be performed with other emotional characters. The performers were instructed to play the given melody in order to render the performance with different emotional expressions, namely 'happy', 'sad', 'angry', 'fearful', 'tender', 'solemn', as well as 'no expression'. The performances were recorded and played to listeners who were instructed to judge all performances with regard to their 'happiness', 'sadness', 'anger', 'tenderness', 'expressiveness', 'fear' and 'solemnity'. The judgements were made on a scale from 10 to 0, where 10 designated maximum and 0 minimum of the respective attribute. The melodies used in this study were either

		existing compositions, or composed specifically for the study.
	Gabrielsson and Lindström, 2001:227	The authors mention a study conducted by Thompson and Robitaille in which composers were asked to write short monophonic melodies to express the emotions of joy, sadness, excitement, dullness, anger and peace. The scores were then transformed to music using computer software. Listeners were instructed to rate the pieces on corresponding emotion scales.
3.2.2 Emotional experiences a) Self-generated	Sloboda and Juslin, 2001:74, 84	The authors referenced a study conducted by Pike. In his study, musically untrained participants were instructed to listen to different pieces of music, and then write down their emotional responses.
	Tan and Kelly, 2004	Groups of formally and informally trained musicians were asked to listen to various pieces of music, and then visually describe their experience of the music using a self-made drawing. They then had to provide a written explanation of their drawing.
3.2.2 Emotional experiences b) Predetermined	Balkwill and Thompson, 1999	Hindustani raga samples were used. Each sample was intended to convey an emotion; therefore the emotion the listeners were supposed to sense was predetermined.
	Balkwill *et al.*, 2004	Listeners were asked to judge the intensity of prescribed emotions (joy, anger or sadness) in excerpts from Japanese, Western and Hindustani music.
	Bresin and Friberg, 2000	Participants were instructed to identify emotional expression in musical examples by choosing from a list of seven predetermined emotions.

	Oh, 2006	Children were asked to draw their emotional responses to musical stimuli. The stimuli were short extracts from classical music pieces that were supposed to correspond to the emotions of happiness, sadness, anger and fear.
	Gabrielsson and Juslin, 1996	Listeners were instructed to judge recordings with regard to their 'happiness', 'sadness', 'anger', 'tenderness', 'expressiveness', 'fear' and 'solemnity'. The judgements were made on a scale from 10 to 0, where 10 designated maximum and 0 minimum, of the respective attribute.
	Becker, 2001:143	Becker quotes the autobiography of Ravi Shankar (1968) who expresses the emotional content of his own musical performance, as determined by the raga he is performing.

2.3 MAP AND STUDIES USED IN THE CURRENT RESEARCH PROJECT

Sections 2.3 - 2.5 (this section of the report and the following two) refer only to the studies and measuring instruments on which the *current* study was based. Each study and measuring instrument is discussed. Figures 2.10 – 2.12 represent the method that was followed in the current study. The numbers in this map correspond with the numbers on the map in the previous section in order for the reader to see which 'methodological path' was chosen for the current study.

Figure 2.10 Map of the current study – profiling

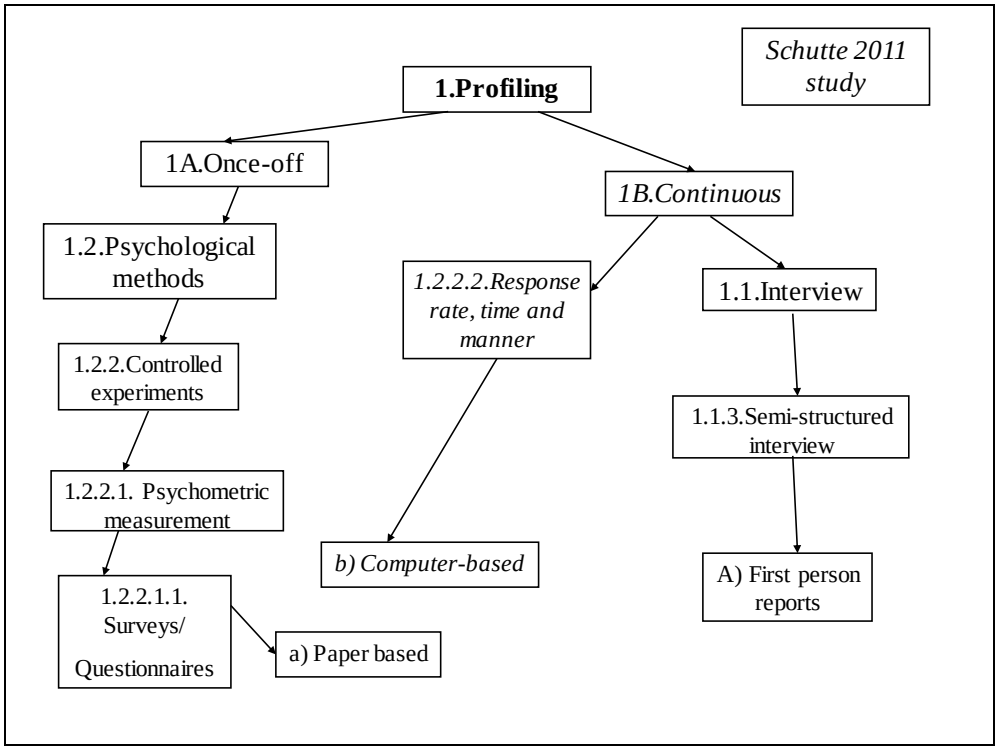


Figure 2.11 Map of the current study – measurement of emotional response

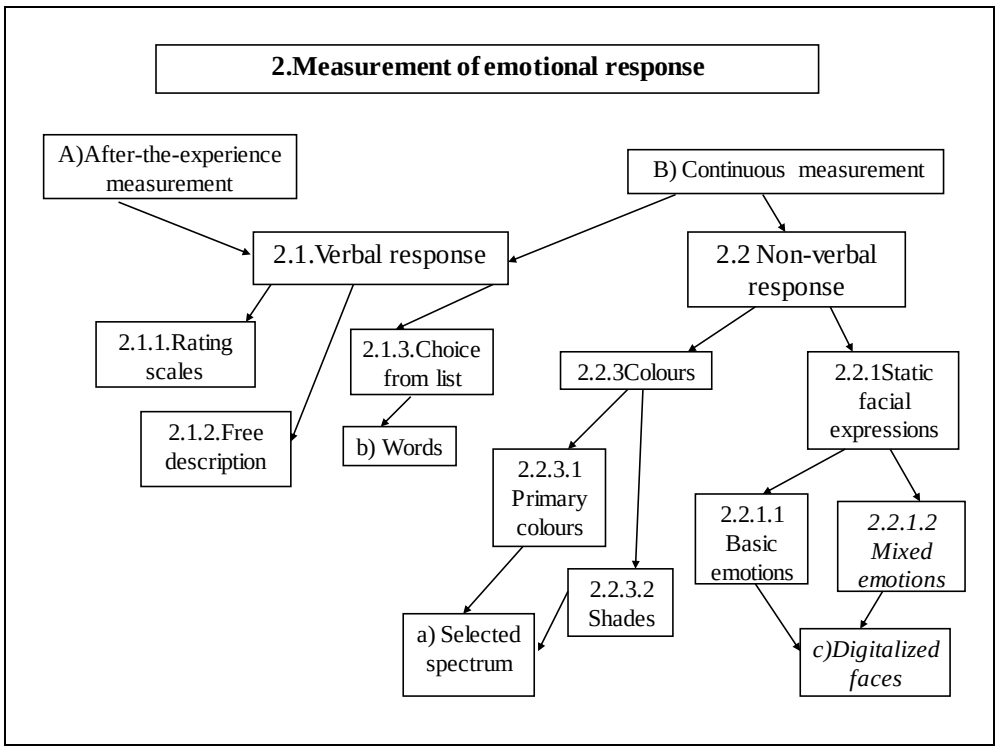


Figure 2.12 Map of the current study – musical stimuli

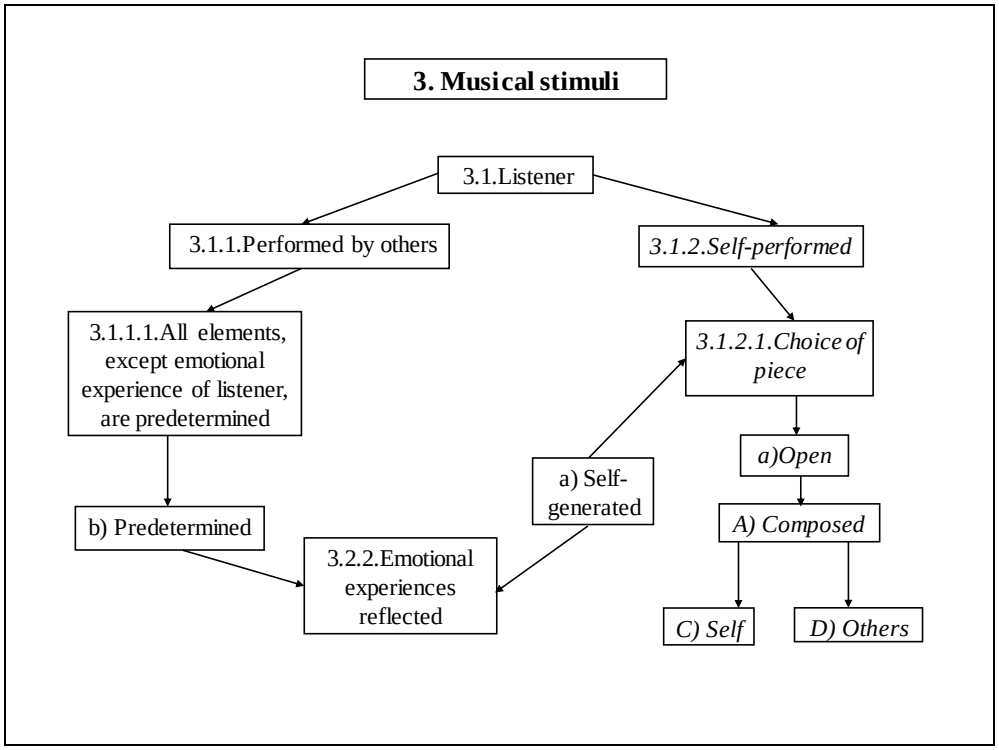


Table 2.3 is a summary of the studies upon which the current study was based. The table also shows where the study fits within the method as it is described in Chapter 3. The table is followed by a discussion of each study.

Table 2.3 Summary of studies used in current study

Schutte 2011 study: map reference (Figure 2.10-2.12)	**Where in method?**	**Study**
	Previous study	Steinberg, 2006
1. Profiling 1B Continuous 1.2.2.2 Response rate, time and manner b) Computer based and 2. Measurement of emotional response B) Continuous measurement 2.1 Verbal response	Ponto Vista, Question 1, (word sorting) and Question 2 (word checklist)	Shaver *et al.*, 1987

2.1.3 Choice from list b) Words		
2. Measurement of emotional response B) Continuous measurement 2.2 Non-verbal response 2.2.3 Colours 2.2.3.1 Primary colours 2.2.3.2 Shades a) Selected spectrum	Ponto Vista, Question 2, colours checklist	Existing research
2. Measurement of emotional response B) Continuous measurement 2.2 Non-verbal response 2.2.1 Static facial expressions 2.2.1.1 Basic emotions 2.2.1.2 Mixed emotions c) Digitalised faces	Ponto Vista, Question 2, facial expressions checklist	Vanger *et al.*, 1998
2. Measurement of emotional response A) After-the-experience 2.1 Verbal response 2.1.1 Rating scales	Ponto Vista, Question 4, rating scales	Balkwill and Thompson, 1999 and Shaver *et al.*, 1987
1. Profiling 1B. Continuous 1.1 Interviews 1.1.3 Semi-structured interview A) First person reports	Semi-structured interviews	Green, 2002

2.3.1 Previous study: Steinberg, 2006

Development

As mentioned in Chapter 1, the current research extends a previous study (Steinberg, 2006) which explored the dynamics of the emotional responses of listeners from two different South-African cultural groups to Hindustani raga samples in terms of overlapping features in their emotional experience of music. The methodology of the previous study (Steinberg, 2006) was based upon a research project by Balkwill and Thompson (1999), as well as a summary of research methods and the Hevner word circle as described by Gabrielsson and Lindström (2001).

Properties, application and interpretation

Listeners described their emotional experiences to two *alap* samples and two *alap-jor* samples that were intended to each convey a different emotion. They did so by providing free descriptions, selecting words from a list, and providing ratings on a five-point scale for psychophysical variables (pitch range, tempo, volume, rhythmic and melodic complexity), as well as by rating the appropriateness of a term to best describe the sample. A follow-up study was conducted which included African listeners and results were presented at the SASRIM 2007 conference (Steinberg and Taljaard, 2007).

When comparing the listeners' experiences, it became clear that the listeners were not able to identify the intended emotion of the samples, except for sample 2. Listeners did, however, generally experienced the emotional content of the samples in the same way, and also experienced the samples as being similar to each other on some level. For both the first and the second sample, judgment of emotion seemed to be related to judgment of psychophysical variables. For sample 3, during the first *alap-jor* sample, there seemed to be a shift in all of the listeners' experiences. Only some of the listeners credited this shift to the entrance of a third instrument, the tabla, approximately in the last minute of the four minute sample, but it could not be determined whether this was the case for all of the listeners since the methodology was not designed to test for it. Listeners also had similar experiences for sample 3; the results also showed some similarities to their experience of the first two samples.

Although the last sample was also an *alap-jor* sample and the tabla entered about halfway through the sample, there did not seem to be such a big shift in listeners' experiences as in sample 3. There were also some similarities in the listeners' experiences as well as similarities in their experiences of the other samples.

When comparing an individual's own experiences of the four samples to each other, one can observe some similarities, but also many differences. Some listeners appeared to be consistent in their experiences of all the samples, while others appeared to have more varied experiences. Questions about the influence of culture, musical training background, listening ability and a listener's personality on experiences of music seemed to be left unanswered and it became clear that the methodology was not sufficiently designed to test for it. The results suggested that listeners might have similar emotional experiences of music, but that these experiences reflect innate and learned characteristics, in which internal as well as external influences can play a role.

Relevance for the empirical study

An evaluation of the methodology used in the previous study (Steinberg, 2006) revealed the following adequacies and inadequacies.

Adequacies:

1) The research design enabled the researcher to test three interpretations of the same listening experiences. The congruence between the results of the three questions was important to the research method, since it suggested that what was being tested, could indeed be tested.
2) There was a strong dependence on the formats of *self-reported* response, which is most commonly used in this kind of research, and proved to be more effective than response judgements by others.
3) The emotional experiences of music that was previously unknown to the listeners, namely Hindustani music, were studied. The use of unknown music gave the researcher some structure against which the data could be interpreted.

Inadequacies:

1) The samples sounded very similar to each other in terms of melodic properties.
2) The choice of samples was, for instance, not made in accordance with the Hevner circle.

3) The samples that contained a partial *jor* section were also perhaps not a good choice. Some participants reacted 'wrongfully' to those samples, which were not supposed to happen.
4) No pre-tests (e.g. personality tests) or post-tests (e.g. interviews) were conducted, which means that certain aspects of the results could only be speculated upon. The lack of profiling was an important shortcoming.
5) Responses were not registered in synchronization with the samples. Some of the listeners answered the questionnaires backwards. Therefore the shift in their experiences in sample 3, for instance, cannot with certainty be credited to the entrance of the tabla. Also, there was no way for the researcher to determine which word was chosen first, in what order the words were chosen, and in relation to which moment in the music they were chosen. This lack of continuous measurement was another important shortcoming.
6) The congruence would have been more convincing if more samples for each raga were used and if they were re-ordered for each question, as Balkwill and Thompson did.
7) The Hevner circle from which the choice list of words were derived, presented some problems. It is an old model, developed in the 1930's, which uses an older form of English. The language differences between cultural periods and groups were apparent when it became clear that some of the words from the list were misinterpreted.
8) The design was exclusively based on written verbal responses, with no opportunity for non-verbal responses.
9) Experiences were reported only *after* participants had already heard the music once.
10) These listening experiences were only recorded once for each participant. There was no longitudinal measurement.

In light of the findings of the study, it was argued that, in order to conduct a more effective study on the emotional content of listening experiences, the methodology needed some improvements. This argument was already referred to in Chapter 1. Among several changes, the most important was the incorporation of continuous measurement. There also seemed to be a need to develop listening profiles of

participants by means of interviews and existing surveys to aid the gathering and interpretation of data.

The following points are similarities between the current study and the previous study:

- A demographic questionnaire is used to obtain personal information about the participant.
- Several interpretations of the same listening experience are an advantage; therefore various verbal and non-verbal response formats are used.
- Self-report response formats are still employed to obtain data during the listening experience.
- The emotional listening experiences to Hindustani music are studied again; this time with four different musical samples for the same prescribed emotion.

Improvements that have been made to the methodology for the current study:

- Even though music from the Hindustani culture is employed, participants also provide responses to familiar or personal musical works, which, in this case, are performed by the participants themselves.
- The emotional experience of music is continuously measured instead of 'once-off' or with an 'after-the-fact'-approach through a newly-developed computer program (Ponto Vista), which keeps track of participants' choices, changes they make to answers, as well as the amount of time they spend on each question. The computer program also keeps track of the music as it unfolds in time. By computerising the questionnaire the emotional responses to listening experiences can be recorded and interpreted with more accuracy.
- The Hevner circle is replaced with Shaver's list of emotion words (Shaver *et al.*, 1987), which is empirically designed, a newer model, and possesses properties that are more suited to the current generation of participants.
- Participants have more freedom of choices than before. For example, they may order and group the words from Shaver's list into their own unique categories.
- Both verbal and non-verbal responses to music are measured by providing a colours checklist and facial expressions checklist to the participants, in addition to the word checklist.

- Participants are tested over a period of time (three test periods), and not just once-off.
- Profiles of participants are constructed through an extended demographic questionnaire, a personality test and a listening test. Interviews are also conducted to supplement the participants' responses during the listening experiences.

2.3.2 Ponto Vista Questions 1 and 2: Shaver *et al.*, 1987

The article written by Shaver and his colleagues (Shaver *et al.*, 1987) was used as the basis for Question 1, word sorting, and for the word checklist in Question 2 of the methodology. In their article they report on two studies, one exploring the hierarchical organisation of emotion concepts, and one specifying the prototypes of five basic emotions. Their report also shows how the prototype approach might be used in the future to investigate the processing of information on emotional experiences and the development of emotion knowledge (Shaver *et al.*, 1987:1061).

These two studies were conducted on the premise of research done on natural categories which suggest a framework for conceptualizing people's knowledge about emotion. Categories of emotions are formed as a result of repeated experiences and become organized around prototypes. Concepts like love, joy, anger, sadness, fear and surprise can be found at the basic level of the emotion hierarchy, and can be applied in the day-to-day distinctions among emotions. It can also be used to process information about emotional events. These concepts intersect with examples people mention when asked to name emotions, as well as with the emotions children learn to name first. It also overlaps with what is known as basic or primary emotions in theories on emotion (Shaver *et al.*, 1987:1061).

In the article, the authors mention several concepts regarding people's knowledge of emotion, their common ability to identify them, to agree on typical antecedents of common emotions, as well as their agreement on the similarity of a diverse array of emotions (Shaver *et al.*, 1987:1061). The article further mentions research suggesting that the various components of emotion knowledge are most likely part of an organised whole, and research based on the notion that most members of a category have shared features which can occupy central places in an organised mental

representation (Shaver *et al.*, 1987:1061). The role of these representations in social interactions is also mentioned (Shaver *et al.*, 1987:1062). The article proceeds to discuss prototype research, as well as the benefits of basic-level categories in emotion knowledge (Shaver *et al.*, 1987:1062-1063). The notion of emotion and emotion categories as fuzzy sets as well as the possibility of analysing these sets from a prototype perspective is explored (Shaver *et al.*, 1987:1063).

The notion that emotion knowledge can be hierarchically represented is explored through the two studies described in the article. The authors attempted to learn more about the emotion prototypes people construct based on the assumption that descriptively rich prototypes will prove useful in other research on emotion knowledge (Shaver *et al.*, 1987:1063).

Question 1, word sorting and the word checklist as used in the current study is based on Study 1 described in the article by Shaver and his colleagues (Shaver *et al.*, 1987:1064-1072). Study 1 explores the hierarchical organisation of the emotion domain and compares it with multidimensional structures identified repeatedly by factor-analytic and multi-dimensional-scaling techniques. Prototype analysis of people's knowledge of a domain consists of two parts: 1) a description of the hierarchical structure of the domain's categories and 2) specification of category prototypes. The goal was to show that a hierarchical representation is intuitively reasonable, related to everyday emotion categorisation, compatible with prototype theory, and informative (Shaver *et al.*, 1987:1064).

Included in the first phase of Study 1 was a sizable number of emotion names that participants could agree were representative of the category *emotion*. English speakers around the world have had little trouble agreeing that a particular psychological-state name designates a relatively good or a relatively poor example of the emotion category. A list of 213 emotion names was compiled. Students in introductory psychology courses rated the emotion names for prototypicality of the emotion domain (Shaver *et al.*, 1987:1064). Participants of Study 1 rated the states named by each of these 213 terms on a 4-point scale from (1) *I definitely would not call this an emotion*, to (4) *I would definitely call this an emotion*. Mean prototypicality ratings were computed and used to select 135 'good' examples of the emotion domain. These

examples were rated highest in the 4-point scale, with one exception; *surprise* was included since theorists have so often designated it a basic emotion (Shaver *et al.*, 1987:1065).

The next phase of Study 1 consisted of similarity sorting. A different group of participants were tested for this phase. During this part of Study 1, the 135 terms were printed on small cards and presented to the participants with the instruction to sort these cards into categories representing their best judgements about which emotions are similar to each other and which are different from each other. Participants were allowed to make as few or as many categories as they liked, which could contain any number of the 135 terms. The number of categories ranged from two to 64. Category size ranged from 1 to 90 terms in a category (Shaver *et al.*, 1987:1065).

The results showed various subordinate clusters, with some of them combined to form larger clusters. There are five or six separate clusters that might be considered basic. The names of these larger clusters are love, joy, anger, sadness and fear. Surprise was also included, although the authors suggested that it may not qualify for basic-level status (Shaver et al., 1987:1065). The results further indicated that, above the basic level, the only other meaningful distinction was between positive and negative emotions. The cluster-analytic results provided three sets of candidates for basicness: a two-term list at a high level of abstraction (positive vs. negative emotions), a 5-6 term list (love, joy, surprise, anger, sadness, and fear), and a 25-term list which consisted of sub-cluster names that were empirically selected, like affection or cheerfulness. The sub-cluster names are indicated with asterisks on the figure below (2.13). The authors decided to accept the 5-6 term list as indicative of basic emotion, since it conforms most closely to basicness criteria (Shaver *et al.*, 1987:1068).

Figure 2.13 Results of the Shaver study

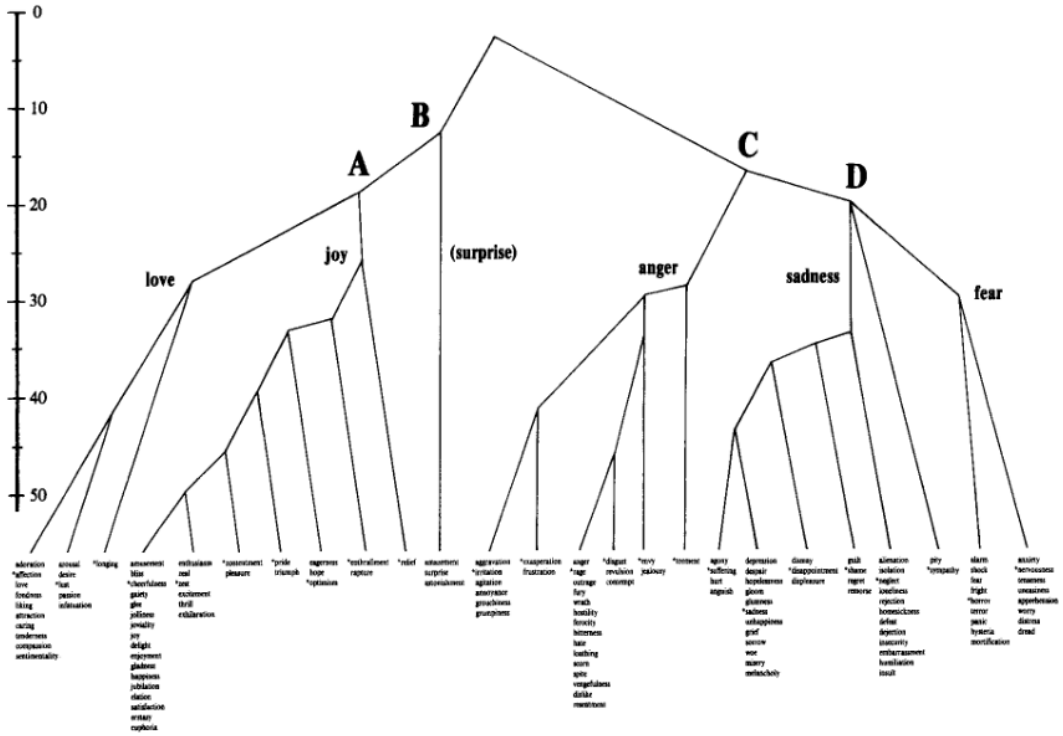


Figure 1. Results of a hierarchical cluster analysis of 135 emotion names. (Cluster strength can be determined by referring to numerical scale at left. Asterisks indicate empirically selected subcluster names.)

(Shaver *et al.*, 1987:1067)

The authors concluded that all of the terms in the emotion lexicon (those that are most prototypical of the category *emotion*) refer to only a few basic-level emotions. Each term seems to specify either the intensity of the basic emotion in question, or the antecedent context in which the emotion arises. The emotion lexicon therefore may refer to the particular basic emotion in question and its intensity, specific eliciting context, or both. In this case, the structure of emotion knowledge doesn't seem as complex as initially expected. The three-dimensional representation of data from Study 1 was also statistically justifiable and highly interpretable. Similarity judgements are therefore proven to be compatible with both hierarchical and dimensional representations. The two forms of representation also reveal different aspects of emotion knowledge (Shaver *et al.*, 1987:1072).

Relevance for the empirical study

The article written by Shaver and his colleagues as described above was selected as the basis for Question 1, as well as for the word checklist in Question 2. It replaces the Hevner circle used in the previous studies (Steinberg, 2006; Steinberg and Taljaard 2007). The following reasons relate to choosing the Shaver study:

- The study was conducted in 1987, which means that it is a newer study than Hevner's, which was conducted in 1936. There seemed to be fewer linguistic challenges when using the newer study.
- Shaver's study was empirically designed and is supported by other research conducted on the subject matter.
- The method used in Shaver's study enabled participants to construct their own categories, which provided each participant with an individual context for the word checklist. A participant's sorting of emotion words provides the researcher with an understanding of his/her emotional world.

2.3.3 Ponto Vista, Question 2, Colours checklist: existing research

The colours checklist presented in Question 2 of the methodology was not based on a specific study, because a suitable study could not be found at the time. Mostly studies that investigated the relationship between colour and emotion were consulted. In a study conducted by Kaya and Epps (2004), college students were asked to indicate their emotional responses to red, yellow, green, blue, purple, black, white, grey, and shades of the primary colours. Another study investigated students' choice of favourite colour, the colour they wore most, and the emotional responses they conveyed to the provided colours. Colours used in this study were white, pink, red, yellow, blue, purple, green, brown, black and grey (Hemphill, 1996:275-276). Colours used in a colour-emotion matching task were red, yellow, dark blue, bright green, brown and black. Subjects were asked to match these colours with facial expressions of happiness, sadness and anger. For a colour preference task, bright blue, dark green and pink were added to the list (Zentner, 2001:390-392).

A study conducted by Da Pos and Green-Armytage (2007) investigated the relationship between facial expressions, colours and basic emotion. The six basic emotions conveyed by the facial expressions used in their study were anger, surprise, disgust, sadness, happiness and fear. Although the colours used in their study were not directly specified in the article, one can observe from the figures provided in the article that they used red, orange, yellow, green, blue, indigo, violet, black, white and grey, as well as various shades of these colours.

The studies mentioned above either did not specify the exact colours used in the studies, or used a colour model that was unfamiliar to the researcher. For these reasons it was decided to use the RGB colour space, so the colours used in the current study could be specified.

The RGB colour space (Red, Green, Blue) can be described as an additive colour space that is based on the RGB colour model. It is defined by the three chromaticities of the red, green, and blue additive primaries, and can produce any chromaticity defined by those primary colours. The complete specification of an RGB colour space also requires a white point chromaticity and a gamma correction curve. It describes what kind of light needs to be radiated to produce a given colour. Light is added together to create form from out of the darkness. RGB stores individual values for red, green and blue (Wikipedia, 2008).

RGB is an especially convenient colour model for computer graphics. The most commonly used RGB colour spaces are sRGB (standard RGB) and Adobe RGB (which has a significantly larger gamut). sRGB was created by Hewlett-Packard and Microsoft Corporation for use on the Internet and World Wide Web, and has been used most frequently since 2007. It is mainly used in consumer grade digital cameras, HD video cameras, computer monitors and HDTVs, because it is considered acceptable for most consumer applications. This means that an image does not need to be converted before being displayed, since all consumer devices use the same colour space. It uses a gamma of 2.2 (Wikipedia, 2008).

Gamut can be defined as a complete subset of colours, and is that portion of the colour space that can be represented or reproduced. The most common usage refers to the subset of colours which can be accurately represented in a given circumstance, such as within a given colour space or by a certain output device. The gamut of a colour is determined in the hue-saturation plane, since most systems are able to produce colours with a wide range of intensity within their colour gamut. Chromaticity can be defined as the quality of a colour, irrespective of its luminance and hue (Wikipedia, 2008).

Relevance for the empirical study

Although the colours checklist presented in Question 2 of the methodology is not based on any specific study, there seems to be enough evidence in research to support the colours and shades used in this study, as can be concluded from the investigations described above. There clearly is a connection between colour and emotion. Evidence of studies conducted on the relationship between colours, facial expressions and basic emotion (e.g. Da Pos and Green-Armytage, 2007) was mentioned before. The six basic emotions conveyed by the facial expressions used in their study were anger, surprise, disgust, sadness, happiness and fear. This corresponds with Vanger's facial expressions, and also overlaps with the basic emotions identified by Shaver and his colleagues. The colours for this study were selected with the aim of providing a scope of choices that can be related to the basic emotions.

The colours checklist in Question 2 consists of 27 colours and is presented to the participant *without* colour names or RGB-values. The list of colours consists of the seven basic colours of the rainbow (red, orange, yellow, green, blue, indigo and violet), together with additional colours, (black, brown, grey, and white). Shades of the rainbow colours as well as a shade of brown and two more shades of grey were added to broaden the participants' scope of choice. Please see Chapter 3 for the specific RGB-values of the colours.

2.3.4 Ponto Vista, Question 2, facial expressions checklist: Vanger *et al.*, 1998

The facial expressions checklist presented to the participants in Question 2 is based on an article written by Phillippos Vanger, Robert Hoenlinger, and Hermann Haken (Vanger *et al.*, 1998). In their article they presented a method of producing prototypical facial expressions of different emotions based on computation and deformation of digitalized facial images. Facial expressions of six basic emotions, namely joy, sadness, fear, surprise, disgust and anger, were portrayed by ten individuals. Each individual facial image was then deformed so as to accommodate to a 'face stencil' defined by standard points on the facial structure. Prototypes for the expressions of each emotion were created by averaging the images of all individual faces, which were reduced to a single computer generated face while holding the facial expression. This procedure was followed for all six facial expressions. A neutral facial expression was also included. Further combinations of upper and lower face parts produced various facial expressions with less clear emotional meaning (Vanger *et al.*, 1998:25).

The authors described the information that a facial expression could convey in terms of a signal system, and remarked that much research had been conducted on rapid facial signals and their role in interpersonal communication, communication of various emotional states in social interaction, as well as the correspondence of expressions to basic emotions. Rapid facial signals can be described as changes in the neuromuscular activity that may lead to visually detectable changes in facial appearance (Vanger *et al.*, 1998:25).

Their article stated that amid this line of research a number of decoding studies have been conducted using facial material of spontaneously produced or posed activity. Ekman and Friesen developed many such systems for the Facial Action Coding System (FACS) manual, usually in the form of photographic, film or video material of real individuals. Different decoding studies have been conducted by various research groups that have developed their own facial material according to the needs of their studies. Consequently, there is a great variability in physiognomic characteristics of

the real individual involved in the production of facial material (Vanger *et al.*, 1998:26).

The authors continued to describe the role of facial structure in facial perception and judgment, and indicate evidence proving that facial structure may influence the way an individual will be judged and responded to. Other studies are mentioned where rounder facial structure was more positively rated than quadratic shaped faces, regardless of their facial expression. They concluded that individual characteristics of facial material needed to be minimised in order to investigate the perceptual and judgmental process regarding facial expressions. Artificially constructed material with the use of computer technology therefore seemed to be a valid alternative to photographic portrayals of facial expressions (Vanger *et al.*, 1998:26-27).

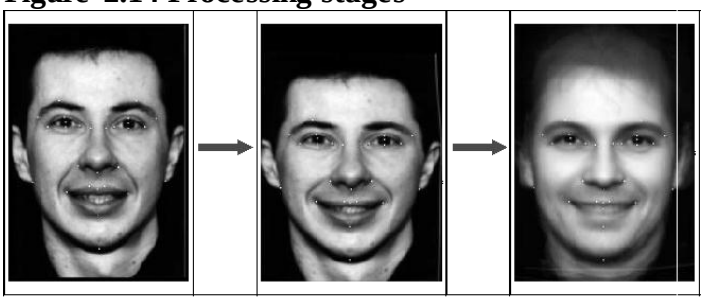
The method Vanger and his colleagues used in their study deals with the creation of artificial facial material depicting emotional expressions with standardised facial morphology (Vanger *et al.*, 1998:30). Ten participants were instructed by a trained FACS coder to voluntarily produce the combinations of action units specified for the six emotions. They were then photographed. A neutral, relaxed face was also photographed. Each photograph was digitalised and stored (Vanger *et al.*, 1998:31). An action unit is the movement as described by FACS which includes all possible movements of the facial musculature that produce a visible change in the face. A combination of different action units results in a facial expression (Vanger *et al.*, 1998:28).

According to Vanger et al. previous studies conducted to create an average face, the pupils of the eyes were used as stable points for photographically blending different faces. In order to produce a face prototype, all individual facial images depicting the same emotional expressions needed to be averaged and blended in terms of their corresponding pixels. This resulted in facial expressions that were blurred and confusing. FACS coding was then not possible, due to the different physiognomical characteristics of the participants' faces. It was evident that averaging procedures based on the centre of the pupils resulted in blurred facial composite images where expression was not portrayed sufficiently (Vanger *et al.*, 1998:31).

For this reason the 'face stencil' was developed, which consists of a greater number of selected points on the face. It allows the reshaping of each face into an averaged facial structure which can be defined into facial composite images. The face stencil was developed in order to overcome the limitations of physiognomic variability so that a standard facial structure could be devised which could be uniform for all facial images (Vanger *et al.*, 1998:31). A standard stencil consisting of 29 points was obtained for each of the six emotional expressions. For each of these 29 reference points an average point was calculated over the corresponding points of the ten original faces of each expression. Next, each individual face was adjusted to the standard stencil of the corresponding emotion in order to construct prototypes. This was done by matching the individual reference points to the standard ones and interpolating all other pixels of the image. Therefore, each individual facial image was distorted to fit the dimensions of the stencil. Each individual face then adopted identical physiognomical dimensions in terms of the length of the nose, distance between the eyes, and so forth. Similar appearances changed on the faces exhibiting the same expression and were defined by identical reference points (Vanger *et al.*, 1998:32).

The figure below (2.14) portrays the processing stages of the facial image. Picture A shows the point referencing of the participant's face; picture B shows the adjustments of the facial image to the stencil, and picture C shows the facial prototype produced by averaging the stencilled facial images depicting the same emotional expression.

Figure 2.14 Processing stages

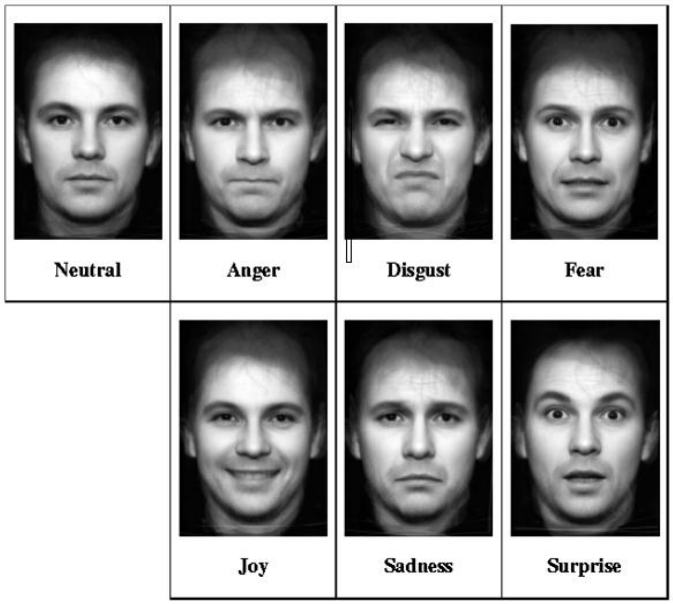


a **b** **c**

(Vanger *et al.*, 1998:32)

Following the same procedure, six facial prototypes of emotions and one neutral were developed for the six emotional expressions (Figure 2.15).

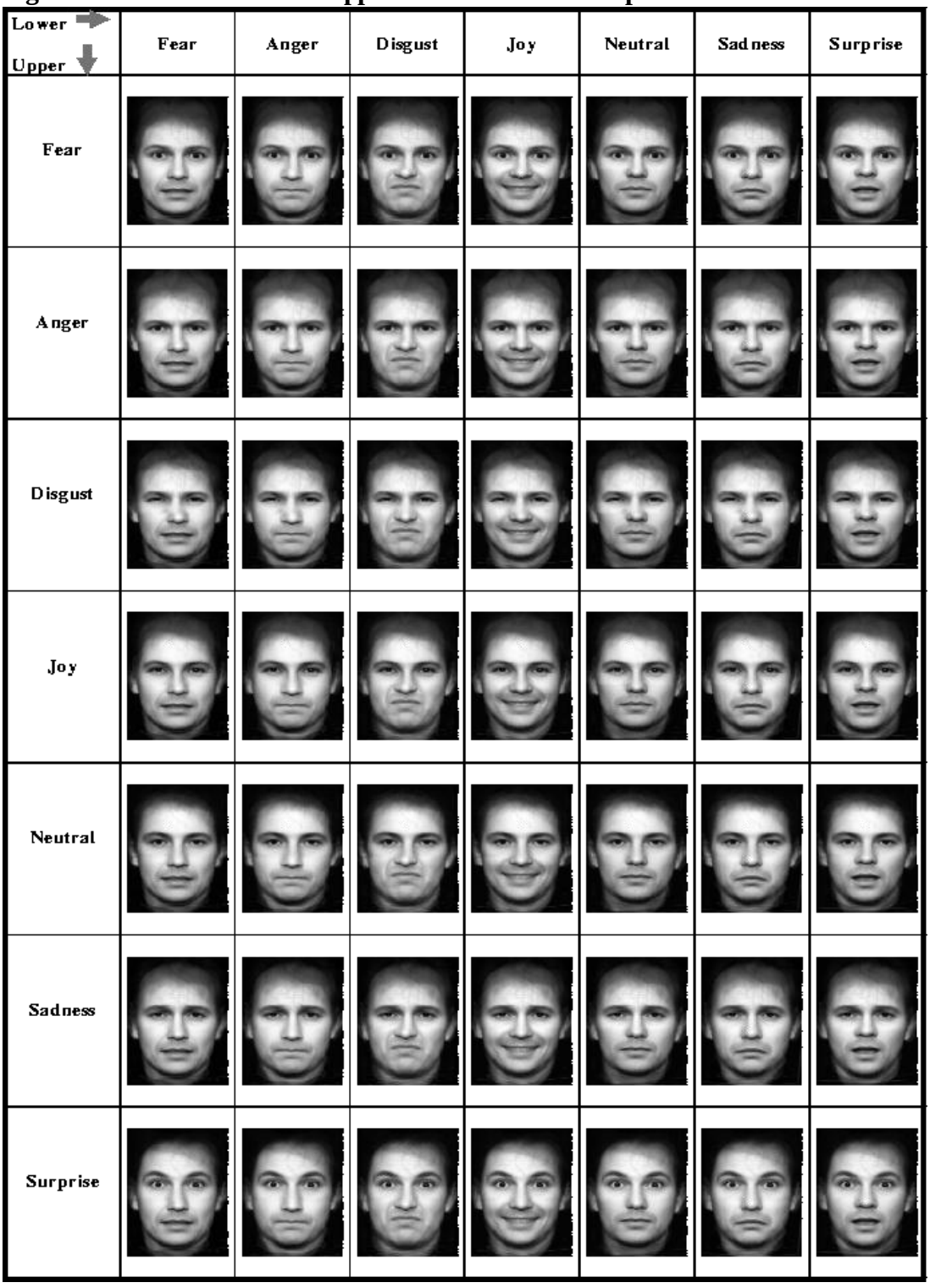
Figure 2.15 Prototypes



(Vanger *et al.*, 1998:33)

These prototypes were coded according to FACS. These seven prototypical expressions of emotion were then further processed with a computer in order to produce other facial expressions with less clear emotional meaning. A number of combined emotional expressions were created by dividing the face in two autonomous areas of upper and lower parts. Forty-nine prototypes of upper and lower face combinations emerged from this procedure (Figure 2.16).

Figure 2.16 Combination of upper and lower facial expressions



(Vanger *et al.*, 1998:33-34)

Relevance for the empirical study

The 49 facial expressions developed by Vanger and his colleagues are presented in Question 2 as a checklist, *without* any indication as to which emotion is being expressed by the combination of the upper and lower parts of the digitalised image. During the listening experiences, the participants can select faces from this list to describe their emotional experiences. The static images of Vanger's study are best suited for the computer program since participants are bound by time. The facial expressions portrayed also correspond to a large extent with Shaver's six prototypical emotions (Shaver: love, joy, surprise, anger, sadness and fear; Vanger: joy, surprise, anger, sadness, fear and disgust).

2.3.5 Ponto Vista, Question 4, rating scales: Balkwill and Thompson, 1999 and Shaver *et al.*, 1987

The article written by Balkwill and Thompson (Balkwill and Thompson, 1999) was used as the basis for Question 4, rating scales, present in the methodology of this study. In their article, Balkwill and Thompson addressed the hypotheses that emotion in music is communicated through a combination of universal and cultural cues, by using a cross-cultural approach. They investigated the ability of listeners to identify intended emotion in music from an unfamiliar tonal system, as well as their sensitivity to associate perceived changes in psychophysical dimensions with those intended emotions. Western listeners rated the degree of four emotions, joy, sadness, anger and peace in twelve Hindustani raga samples. Each sample was intended to convey one of the four emotions. The listeners also provided ratings of psychophysical variables (tempo, rhythmic complexity, melodic complexity and pitch range). Their findings suggest that the listeners were able to identify musically expressed emotion in an unfamiliar tonal system and that it was facilitated by psychophysical cues (Balkwill and Thompson, 1999:43).

Balkwill and Thompson pointed out that research on emotional meaning in music focused mainly on music from Western tonal systems and the judgments of Western listeners, which meant that there was a very small basis for evaluating the

enculturation hypothesis of the perception of emotion in music. They also suspected that there may be universal influences underlying musically expressed emotion, such as psychophysical dimensions of music (any property of sound that can be perceived independent of musical experience, knowledge or enculturation) (Balkwill and Thompson, 1999:44). They argued that the association of psychophysical dimensions of music with the perception of emotions may also be the result of enculturation and may differ between cultures. If so, listeners should not be able to identify intended emotions in music from an unfamiliar tonal system. However, if those associations are not determined by culture, then listeners should be able to go beyond cultural boundaries and identify intended emotion in music from an unfamiliar tonal system. Their study was specifically designed to investigate this possibility (Balkwill and Thompson, 1999:46).

Studies have shown that a number of psychophysical dimensions have been linked with judgments of emotion in music. This includes tempo, melodic contour, harmonic complexity, melodic complexity, rhythmic complexity, articulation, dynamics, consonance and dissonance, pitch register, and timbre (Balkwill and Thompson, 1999:48). Balkwill and Thompson examined five of these dimensions in their study, namely tempo, melodic complexity, rhythmic complexity, pitch range and timbre (Balkwill and Thompson, 1999:48).

Thirty Western listeners participated in the study, and were asked to rate the degree to which they believed an emotion was conveyed in twelve *alap* portions of Hindustani ragas on a 9-point scale. Next they were asked to rate each piece on similar scales for the other emotions, as well as for the psychophysical variables that have already been mentioned. Definitions and verbal explanations of each of the psychophysical variables were provided, e.g. was the tempo very fast (9) or very slow (1), or e.g. was the melody very simple, meaning they could sing it to someone (1) or would it be difficult to do so (9) (Balkwill and Thompson, 1999:51-52).

Relevance for the empirical study

For the current study, the following psychophysical variables were tested in Question 4: pitch range, dynamic level, tempo, rhythmic complexity and melodic complexity. It overlaps with the variables that Balkwill and Thompson tested. The rating scales for this study was adjusted to 5-point scales, as opposed to Balkwill and Thompson's 9-point scale, the reason for this being that nine options could possibly produce results that were too varied. The six basic emotions (love, joy, anger, sadness, fear and surprise) identified in the study conducted by Shaver were also prepared into a rating scale according to which participants were asked to rate the appropriateness of the six emotions to the music they listened to. This concept also overlaps with the procedure followed by Balkwill and Thompson.

2.3.6 Semi-structured interviews: Green, 2002

The general structure of the interviews was based on the work of Lucy Green (2002). She conducted interviews with informally trained musicians and the questions she asked them were relevant to the current study. The structured questions from Green's work concerned participants' involvement in music and performing, music education of any kind, experiences at school that are related to music, recording background, and their emotional perceptions of their own performances. The rest of the questions that were asked during the interviews were related to the results of Part II of each of the three test periods. The results were discussed individually with each participant in order to clarify his/her answers and inquire into the manner in which he/she has completed that particular session. The questions that were extracted from Lucy Green's work, as well as some additional general questions that the participants had to answer, are presented in Chapter 3.

2.4 MEASURING INSTRUMENTS USED IN THE CURRENT RESEARCH PROJECT

Table 2.4 is a summary of the measuring instruments that were used in the current study. The table also shows where the measuring instrument fits within the method as

described in Chapter 3. The table is followed by a discussion of each measuring instrument.

Table 2.4 Summary of measuring instruments used in the current study

Schutte 2011 study: map reference (Figure 2.10-2.12)	**Where in the method?**	**Measuring instrument**
1. Profiling 1A. Once-off 1.2 Psychological methods 1.2.2 Controlled experiments 1.2.2.1 Psychometric measurement 1.2.2.1.1 Surveys/questionnaires a) Paper based	Listening Profiles	Demographic questionnaire
	Listening Profiles: listening test	Tomatis test
	Listening Profiles: personality test	NEO PI-R
2. Measurement of emotional response	Ponto Vista	Ponto Vista

2.4.1 Listening Profiles: Demographic questionnaire

Development

The demographic questionnaire is based upon the demographic questionnaire used in a previous study (Steinberg, 2006:27, 67).

Properties, applications and interpretation

Questions concern a participant's name, age, gender, nationality, cultural background, maternal language, year and course of study as well as preference for musical styles and instruments. Other questions concerning music education, musical preferences, health, and family relations to parents and siblings were added due to the nature of this study. The information provided in this questionnaire serves as part of the

participant's profile, which provides the researcher with a context within which the data of the participant's listening experiences can be interpreted.

Relevance for the empirical study

The demographic questionnaire serves as a tool to establish an overall historical profile of the participants. Please see the Addendum for the complete demographic questionnaire.

2.4.2 Listening Profiles: Tomatis test

Development

Alfred A. Tomatis was a French physician and ear-nose-throat specialist. In the late 1940s, Tomatis undertook a series of experiments in order to determine the effect of noise upon the hearing ability of humans. The results of those experiments convinced him that the vocal and musical abilities of individuals are directly linked to their ability to hear. He also had a clinical practice where he could apply his experimental findings therapeutically. He discovered the far reaching effects of sound stimulation, used over a sufficient amount of time, to condition one's ear to improve discrimination between and analysis of the sounds of music and voice. Starting in the early 1950s, Tomatis developed an electronic device designed to 'train the ear' which can improve both voice production and musical skill. This device, called the 'Electronic Ear' has also been used to improve language, learning and communication skills. The technology, program, and procedures that Tomatis developed can be described as 'audio-psycho-phonology'. It is also known as the 'Tomatis Method'. It indicates the interaction between an individual's listening and hearing potential, psychological attitudes, and control of speech and language. It specifically stimulates the rich interconnections between the ear and the nervous system. It is a non-invasive sound stimulation program which is used to enhance listening skills, which in turn contribute to the enhancement of other skills and help to overcome listening-related problems. Guidance counselling is also provided as part of the program (Madaule,

1997:1, 2; Thompson and Andrews, 2000:175; Van Jaarsveld, 1974:121; Vercueil, 2010:41).

Tomatis identified several ‘laws’ that govern the ear and voice and their effects on a person’s behaviour and abilities. The principle law with its two outcomes form the scientific basis for the Method and initially drove the fundamental engineering concept behind the development of Tomatis’s major invention, the Electronic Ear. It was developed in order to assist the human ear to establish or re-establish its full potential. The three ‘Tomatis Laws’ can be summarised as follows: 1) the voice can only produce what the ear is capable of hearing, 2) auditory capability can be restored by re-educating the ear, which will restore the voice as well, and 3) with sufficient controlled auditory stimulation - listening to one’s own voice - the changes in one’s self-listening behaviour and phonation will be maintained (Thompson and Andrews, 2000:176; Vercueil, 2010:42, 49).

Other important findings of Tomatis are:

- Physical or psychological causes might influence listening and/or hearing capabilities.
- An increase of 6dB. per octave in the middle-frequency region is necessary for the correct identification and reproduction of music.

(Van Jaarsveld, 1974:138-139; Vercueil, 2010:52)

Tomatis was convinced that the ear plays an integral role in relationships, since he considered the ear to be the “door to communication with the world of sound”. This “door” could be open or closed to communication. When closed, any meaning of the world of sound is lost, but when open, a whole new dimension appears (Van Jaarsveld, 1974:134).

Hearing can be defined as a passive, automatic, sensory act, while listening is an active, voluntary act, which is a shift from awareness to perception, which also includes the ability to analyse sound with speed and precision. It is possible for a

person to have good hearing, but poor listening. Listening affects communication, motivation, rhythm and coordination, as well as learning potential. Good self-listening is also required for the recognition of the quality of one's communication, the evaluation of speech volume and intensity as well as the continuity of speech. Poor listening is often the result of emotional or physical trauma, which leads to disharmony in one's total integration. Symptoms include problems in areas of attention span, learning, motor control, receptive and expressive language, self-esteem, behaviour and attitude (Thompson and Andrews, 2000:176; Vercueil, 2010:56, 57).

The Tomatis Method consists of an initial assessment, followed by listening sessions designed to train the ear to restore its full listening potential (Vercueil, 2010:57). The initial assessment consists of the Tomatis listening test which is an auditory perception test that combines information on hearing and the utilisation of the desire to listen. It serves as a diagnostic tool that provides information regarding the use of listening potential, body-related functions such as posture, muscle tension, balance, motor functions, and energy levels. The test also consists of a psychological dimension (Vercueil, 2010:90). A modified tonal audiometer is used to assess the auditory threshold curves, both of air and bone conduction. Tonal differentiation, which is the ability to discriminate between various sound tones, is also evaluated. Other elements of testing include a lateral dominance checklist and a drawing test. Close observation of the individual's listening style and communication behaviour is made throughout (Madaule, 1997:2).

The Tomatis listening test shows individuals' listening strengths and weaknesses as well as distortions in relation to the ideal auditory curve and reveals impaired listening, which provides insight into the way someone relates to the environment. It also serves as a tool to observe changes that might occur during and after exposure to the listening program (Vercueil, 2010:58). It consists of four evaluations which measure curves of the threshold in air and bone conduction of sounds, auditory spatial orientation (directivity), auditory selectivity and auditory laterality. The results are indicated on a listening diagram. Peaks and dips in the hearing curve provide

diagnostic possibilities, but all parameters of both ears should be taken into account when analysing the data. The curves of the one ear can be compared with the curves of the other. Any distortions in the curves of the left ear that are repeated in the curves of the right ear point towards an affection that is still being experienced and might be chronic in nature (Van Jaarsveld, 1974:234).

According to Tomatis's observations, the right ear coordinates phonation and musical capabilities, and therefore an excellent performance depends on right-ear dominance. Right-ear dominance appears in people with long and intensive musical training. This capability is known as the 'musical ear' and refers to the average auditory curve in the profile of musically gifted and musically appreciative people. This profile is also the profile for good hearing. It does not guarantee excellence; it simply means that a student who has a 'musical ear' is easier to teach. Good auditory lateralisation implies the right ear becomes the 'leading ear'. The term 'leading ear' refers to a person's chosen, preferential ear which focuses on specific sound while the other ear provides an overall image of the sonic background. Failure to develop the right ear as leading ear often results in language and learning problems (Vercueil, 2010:49-52, 55).

Properties, applications and interpretation

The Tomatis Method is used to enhance listening skills, which in turn improves other skills and works on a variety of problems because of the rich interconnections between the ear and the brain stem, the cerebellum, and the higher cortical centres. Thus, problems and deficits that have a neurological basis can be helped by a specific application of sound stimulation with certain characteristics. Some of the problems to which the Tomatis Method has been applied include head injury and other neurological injuries such as strokes, developmental delays, autism, PDD, Multisystem Sensory Disorder and other types of brain damage from birth, attention deficits, and learning disabilities (Thompson and Andrews, 2000:184). Research and experience with the Tomatis Method show that it provides a great benefit for developing control of speech, language, reading and writing, comprehension and perceptual processing, attention, concentration, memory and auditory processing

abilities (Vercueil, 2010:41,64,65; Thompson and Andrews, 2000:184). It also includes positive effects on communication, learning, psychological well-being, musical skills and the reduction of psychological symptoms. It has physiological benefits such as improved hand-eye coordination, balance and motor function, and furthermore is effective in the reduction of stuttering, anxiety, depression and other pathological tendencies (Vercueil, 2010:41, 64-66).

During the Tomatis listening test, auditory spatial orientation (directivity) is evaluated in an individual's ability to identify the direction from which a sound is coming to the extent to which the individual has integrated the notion of space and time in the perception of sounds. Problems regarding directivity are indicated opposite the relative frequencies at the bottom of the listening diagram. This test is done in conjunction with the bone conduction test (Van Jaarsveld, 1974:221). The selectivity test assesses the ability to recognise pitch differences in neighbouring sounds (Van Jaarsveld, 1974:220). If the individual's selectivity is closed at a certain frequency, it implies that the integrating function of all the frequencies above that level will also be affected. For instance, closed selectivity from 2000 Hz indicates that the individual will most likely have trouble to observe sounds correctly, meaning that the individual can still hear, but does not listen or integrate the sound in that frequency or beyond (Van Jaarsveld, 1974:221). Problems regarding selectivity are indicated opposite the relative frequencies at the top of the diagram. The laterality test assesses ear dominance, and the relationship between the ears is indicated by a number written on the diagram (Van Jaarsveld, 1974:222; Vercueil, 2010:91). The air conduction curve is indicated in blue, and the bone conduction curve in red (Van Jaarsveld, 1974:220).

Variances in the curve could have several implications for a musician's abilities. For instance, when a curve is devoid of its treble beyond 2000 Hz, tonal reproduction is still secure, but difficulties in quality control could be experienced (Vercueil, 2010:52). Variations in the curve between 1000 Hz and 2000 Hz lead to faulty intonation while musical receptivity and quality appreciation are reserved. Disjointed curves above 1000 Hz lead to difficulties in the reproduction of music (Vercueil,

2010:53). Dips from 500 Hz to 1000 Hz indicate insensitivity to music. A completely flat or disjointed curve indicates an 'unmusical ear' (Vercueil, 2010:54).

The curve that indicates good hearing can be described as ascending from 500 Hz (C above middle C) to 4000 Hz (C four octaves above middle C) where it forms a dome between 2000 Hz and 4000 Hz, with a slight drop between 4000Hz and 6000 Hz (Van Jaarsveld, 1974:124,126; Vercueil, 2010:51).

The air conduction curve indicates an individual's external relationships. It specifies how the individual listens to other people. The bone conduction curve is connected to the self-listening process and can be described as being representative of the individual's 'inner world' and how he/she listens to his/her own voice. It indicates somatic responses corresponding to different frequencies, providing information regarding posture, especially the spinal column and the head. The relationship between air and bone conduction is important since it reveals social and external behaviour against internal experience and reaction. Ideally, the two curves should be parallel to each other without any distortions, indicating integration in both the milieu and the self. If the two curves cross each other at any point, meaning that the bone conduction curve is then above the air conduction curve, it indicates that the individual has a preference for his/her 'inner world'; he/she is more likely to observe inner sounds and might have trouble with communication. Characteristic of such a person is that he/she is very self-centred and could possibly be egocentric and narcissistic in his/her behaviour (Van Jaarsveld, 1974:233; Vercueil, 2010:92).

Tomatis emphasised that the interpretation of the listening test only represents hypotheses and not proven theories, thus warning against a rigid interpretation. He also emphasised an overall approach in the interpretation of the listening test. It is necessary to interpret all parameters for both ears simultaneously when analysing the listening test (Vercueil, 2010:91-92). According to Tomatis, a listening problem that is not the result of an organic lesion has a psychological origin. Therefore the

symbolism of the left and right ear diagrams and three frequency ranges are also taken into account (Vercueil, 2010:92).

Reliability and validity

The Tomatis listening test was recommended by the Institute for Psychotherapy and Counselling, Potchefstroom. It was conducted by a speech therapist at the Institute, and evaluation reports were provided to the researcher by a clinical psychologist (with a D. Phil as highest qualification) of the Institute.

Relevance for the empirical study

The methodology for this study includes the Tomatis listening test in order to determine the possible presence of dysfunctional hearing and listening ability in the participants. The Tomatis listening test is an auditory perception test that combines information on hearing with the application of the need to listen. Listening strengths and weaknesses are revealed by comparing the air and bone conduction curves of the person who has been tested to the curves of the 'ideal good functioning ear'. This comparison therefore reveals impaired listening (Vercueil, 2010:58). It provided the researcher with an evaluation of each participant's auditory orientation, selectivity, laterality and the thresholds of air and bone conduction.

2.4.3 Listening Profiles: NEO PI-R

Development

The NEO Personality Inventory Revised (NEO PI-R) was developed by Paul T. Costa Jr. (Ph.D.) and Robert R. McCrae (Ph.D.) (Piedmont, 1998:80, 106). It was specifically designed to capture the dimensions underlying the personality model called the Five-Factor Taxonomy (Piedmont, 1998:2). The facets selected by the developers were chosen because of their psychological significance in literature (Piedmont, 1998:41). They endeavoured to create non-redundant scales that still had

the ability to assess the five dimensions overall. They were able to do so with a technique called validimax factoring. It establishes the structure of a scale by its pattern of relationships with external criteria. The facets for each domain were designated on the basis of their being psychologically applicable and descriptively diverse (Piedmont, 1998:35).

Properties, applications and interpretation

The NEO PI-R is a systematic assessment of character, emotional, interpersonal, experimental, attitudinal, coping and motivational styles, as well as levels of personal well-being (Mosupye, 2007; Piedmont, 1998:70). It was designed to access an individual's fundamental temperaments. Scores are more likely to remain constant over time and across situations. When testing personality, it is important to keep in mind that people develop extensive behaviour repertoires for fulfilling their needs. In other words, they are able to adapt their behaviour according to the environment they find themselves in without changing the original need (Piedmont, 1998:4). Therefore, changes in behaviour may not necessarily signal a change in the need being expressed (Piedmont, 1998:5).

The NEO PI-R is not a measure of psychopathology, nor was it designed to capture such types of processes. Therefore, any high or low scores on any of the scales should not be seen as indicative of anything abnormal. The scores need to be interpreted from a non-clinical viewpoint since they speak to the regularities in an individual's on-going behaviour. However, some relationships that define 'characterological' impairment can be expected (Piedmont, 1998:58).

The NEO PI-R uses a five-factor model of personality which represents a comprehensive, empirically based taxonomy of personality traits. Assessment of these five domains is made easier by the fact that each domain consists of specific facet scales. This allows for a more precise and differential evaluation of the individual's position within each personality dimension (Piedmont, 1998:31).

The five-factor model is a comprehensive description of normal personality dispositions capable of organizing individual-difference constructs from a wide

collection of theoretical models. It provides a high degree of empirical 'power' in predicting several life outcomes. The factors are descriptive of non-clinical functioning, and can contribute in various ways to assessing abnormal functioning (Piedmont, 1998:58).

The NEO PI-R is self-administered, available in two versions, and has been translated into many languages. Form S is designed for self-report, while form R is an observer or rater report. Data from form R can be used when data from a self-report cannot be trusted, e.g. in a case where an individual is mentally compromised. It can also be used to provide another perspective to the assessment of an individual, or evaluate groups and their interpersonal dynamics. The rater version consists of the same items, but phrased in the third person (Mosupye, 2007; Piedmont, 1998:35-36).

The NEO PI-R consists of 240 items that an individual answers using a five-point scale with (1) being "*strongly disagree"* and (5) being "*strongly agree"*. The items are comprised of simple sentences which describe specific behaviours or attitudes (Piedmont, 1998:35). It furthermore contains 3 validity items, with an administration time of 35-45 minutes. Individuals taking the test should at least have a Grade 6 reading level (Mosupye, 2007). The NEO PI-R measures the five major domains of a participant's personality and contains six facet scales within each domain which assess more specific aspects of the domain. This presents the researcher with a comprehensive assessment of each participant's normal adult personality (Piedmont, 1998:35).

The domains as well as the facet scales which define each domain are:

1) Neuroticism: Anxiety, Angry hostility, Depression, Self-consciousness, Impulsiveness, Vulnerability
2) Extraversion: Warmth, Gregariousness, Assertiveness, Activity, Excitement-seeking, Positive emotion
3) Openness: Fantasy, Aesthetics, Feelings, Actions, Ideas, Values
4) Agreeableness: Trust, Straightforwardness, Altruism, Compliance, Modesty, Tender-mindedness

5) Conscientiousness: Competence, Order, Dutifulness, Achievement striving, Self-discipline, Deliberation

(Piedmont, 1998:36)

The purpose is to obtain a detailed assessment of normal personality. It can be used in human resource development, industrial and organisational psychology, vocational counselling and clinical practice (Mosupye, 2007). All of the information provided by the NEO PI-R contributes to clinical outcomes indices beyond information obtained only from symptom-based scales. The five-factor model can provide a salient framework for contextualizing an individual's presenting problems. It is therefore also a useful clinical instrument, since it can assist a therapist to organize the information in clinically relevant ways (Piedmont, 1998:70).

The values of personality assessment can be stated in a clinical context where technologies are needed to accelerate the evaluation and diagnostic process of an individual. The NEO PI-R is useful in this context since it helps to identify the types of needs the individual wants to fulfil, interpersonal style, value systems, and levels of on-going emotional distress. It is also helpful in identifying therapeutic interventions that may be effective. It furthermore has the ability to provide a well-documented report in order to determine treatment effectiveness (Piedmont, 1998:9).

Six ways that the NEO PI-R can be employed in a clinical context:

- Understanding the individual (Piedmont, 1998:11).
- Differential diagnosis. The NEO PI-R is useful for providing information relevant to making a diagnosis or for ruling one out (Piedmont, 1998:12).
- Empathy and rapport. The information provided by the assessment can help encourage understanding of and empathy with the individual. The NEO PI-R provides a more holistic, person-centred orientation that captures the individual's growth potential (Piedmont, 1998:12).
- Feedback and insight. By providing the individual with feedback on the results of the NEO PI-R, the individual is given hope and understanding. It can also

be used to create a common framework for the therapist and the individual from where problems, treatment goals, change, and growth can be presented, identified and addressed. It can also assist in establishing meaningful dialogue between the therapist and the client. The report generated for the NEO PI-R is readily accepted by individuals and contain little objectionable material (Piedmont, 1998:12-13).

- Anticipating the course of therapy. The NEO PI-R provides useful information regarding an individual's level of cooperation, motivation to work, and capacity for therapeutic change – characteristics that need to be accommodated in treatment (Piedmont, 1998:13).
- Matching treatments to an individual. Personality measures can facilitate matching the right kind of therapy that will complement the client's style (Piedmont, 1998:13-14).

The empirical basis and conceptual clarity afforded by the NEO PI-R make it ideal in comparing individual differences across sub-disciplines in social sciences. The dimensions of the test are furthermore able to describe both normal and clinical people (Piedmont, 1998:31).

Scoring options include hand scoring, software scoring and bureau service (Mosupye, 2007).

When interpreting the data gathered by the NEO PI-R, the various facet scales and domains should be interpreted in relation to one another rather than in isolation. Any interpretation must include all five domains together with an evaluation of the facet scales (Piedmont, 1998:126-127).

During the interpretation of data, Raw scores as well as T-scores are calculated for each facet scale and domain. The range within which data is interpreted are determined by the T-scores, and can be described as very low, low, average, high and very high. T-scores below 35 are considered to be very low; between 35 and 45 are described as low. When a T-score is above 55, the individual is seen as a high scorer.

T-scores above 65 are considered very high. The higher or lower a score is the more characteristic the descriptions will be of the individual's personality. If an individual have a score between 45 and 55 (in average or mid-range), classification is still possible, but the description will be less accurate. If the other scores are in the high range, then the average score should be considered as low, and vice versa (Piedmont, 1998:116).

In developing the facets, every effort was made to create scales that were as non-overlapping in content as possible, while still remaining on the same domain. By minimizing redundancy among the facets, the test developers maximized the interpretability of each scale (Piedmont, 1998:84). Definitions of each domain and facet scale, as well as descriptions of the five different ranges are provided in Addendum. These definitions assist an evaluator when interpreting the data gathered by the NEO PI-R.

According to Piedmont (1998:79), psychological test interpretation requires of the evaluator a solid grounding in personality theory, an appreciation of the construct validity of the test being used, as well as a firm understanding of psychometrics.

When analysing the results of a NEO PI-R, one should keep in mind that the qualities represented by these scales relate to the normal range of psychological functioning. Scores on the NEO scales reflect the motivations, tendencies, and capacities that characterise the individual's on-going interaction with his environment. Personality in this context refers to the consistent ways in which individuals perceive the universe, their place in it, and the directions in which they want to move. No value judgments are to be imposed when interpreting the scores and an equal view of all qualities should be kept throughout (Piedmont, 1998:57). All personalities are to be treated with fairness since personality assessments inventories are used to assess needs, identify complementing environments and determine how well an individual's personality fits in with the demand of a given situation or environment (Piedmont, 1998:58).

Reliability and validity

Internal consistency and test-retest information for the domains and facet scales are represented in literature, with the alpha reliabilities for the facets ranging from .56 for Tender-mindedness to .81 for Depression; and for the domains from .86 for Agreeableness to .92 for Neuroticism. Other tests have shown the internal consistency for both forms S and R to range from .86-.95 for domain scales and from .56-.90 for facet scales (Mosupye, 2007). The retest reliability values are in the .7 to .8 range, suggesting very good stability over time. A 25-year retest coefficient for the five major personality dimensions shows an 80% of the variance as stable (Piedmont, 1998:36-38). The NEO PI-R was designed to capture genotypic qualities, referring to the basic underlying composition of an individual's nature. These qualities should not vary over time, and is proven by the NEO PI-R's retest capabilities (Piedmont, 1998:38-39). The NEO PI-R was validated against other personality inventories and projective techniques (Mosupye, 2007).

Validity scales have been widely accepted by test users (Piedmont, 1998:79). Much effort has been made in order to develop 'content-free' validity scales for the NEO PI-R. Such measures rely on response styles rather than on item content, meaning that the consistency with which an individual responds to items having similar content is evaluated (Piedmont, 1998:81). With the belief that self-report data are relatively sound, as well as with the support of studies that prove this notion (Piedmont, 1998:83), the developers of the NEO PI-R did not initially include any type of validity scale. The only presence of such a scale is three short items at the end of the NEO PI-R that asks the participant to confirm if he/she has completed all items in the proper way. A 'no' to any of the items would indicate a questionable procedure (Piedmont, 1998:84).

Although it was not specified in the evaluation reports which validity scale was used to verify the data gathered by the NEO PI-R tests conducted for this particular study, the evaluator did in fact test for it. The validity scales of all reports suggested a valid profile for all participants of this study.

Relevance for the empirical study

The NEO PI-R was used in the current study as part of the listening profile (Part I) to help determine the personal background of the participants. It was included in the methodology as part of the improvements that were made, in order to determine whether a participant is experiencing the music used in this study based upon his/her personality, or if the experiences are related to properties in the music itself. For the current study, test form S, designed for self-report, was used. The results were analysed and interpreted by the same clinical psychologist that wrote the evaluation reports for the Tomatis listening test. The NEO PI-R evaluation reports were submitted to the researcher.

2.4.4 Ponto Vista

Development

The questionnaire that the participants had to complete during Part II of the methodology concerning emotional experiences of music was computerised into a program called Ponto Vista. Ponto Vista was developed specifically for this research project in order to capture the emotional responses of a user to various musical tracks in a consistent and continuous way. The creative design was done by the researcher to insure that all questions, instructions and procedures could be followed as determined by the research design. The researcher was assisted by a conceptual designer. Implementation and program design was done by a programmer that specialised in writing source code in the Java Program Language. A logo, desktop icon and loading window was designed by a local graphic design company and added to the program in order for a user to be able to identify the program on any computer.

Properties, applications and interpretation

Since the program is discussed in more detail in Chapter 4, only information relevant to the aims of this chapter will be provided here. The program consists of various windows and navigation controls that can be used during a session. The main window allows the user access to the program, as well as access to a playlist that was also built into the program. The playlist can be edited to determine which musical tracks must be played during the session as well as which interfaces must be used when listening

to a musical track. The program therefore has the ability to create different user profiles that can be customised according to requirements. It furthermore contains an introduction window, where an overview of the questions is provided, which a user must complete when listening to a musical track. There are also visual indicators built into the program to inform a user of his/her progress with the session. Each question of the methodology in Part II together with their requirements was programmed into several windows that a user can access to complete the questions. The program was also designed to prevent a user from continuing should they not comply with the requirements of the question. At the end of the session, the program automatically generates a report which contains the results of the session. It presents every change and sorting a user made, the time a musical track started and ended and all the selections made while using the various interfaces, the complete free description, and the selection of numbers and changes to the rating scales.

'Ponto Vista' is Portuguese and means 'vantage point'. The program used in Part II of the methodology was named Ponto Vista since it was one of the desires of the researcher to provide the participants with an opportunity for self-observation through the program. In other words, to provide the participants with a vantage point that could be observed and measured. The participants were encouraged to use the program to their advantage, to learn from their experiences through the program and apply those experiences in their development as musicians. The intention was for them to use it as a tool to explore the emotional dynamics involved when music is performed and perceived by the same person. The program can be described as a mirror which can allow a person to observe his/her emotional world in a particular way, without outside judgment.

Although the results of the program can be analysed and interpreted in many ways, they were interpreted for this research project in accordance with the goals of the current study.

Relevance for the empirical study

Ponto Vista is a measuring instrument that is applied in Part II of the methodology.

2.5 OTHER 'PATHS' RELATED TO THE CURRENT STUDY

Table 2.5 is a summary of other decisions that were made regarding the methodology of the current study. It is followed by a discussion of each decision. The table shows where the decision fits within the method as described in Chapter 3.

Table 2.5 Summary of methodology-related decisions

Schutte 2011 study: map reference (Figure 2.10-2.12)	**Where in the method?**	**Decision**
	Overall design	Mixed method
2.Measurement of emotional response	Ponto Vista	Three Test Periods vs. one Test Period
2.Measurement of emotional response	Ponto Vista	Continuous measurement
2.Measurement of emotional response	Ponto Vista	Self-report
2.Measurement of emotional response	Ponto Vista	Verbal and non-verbal response formats
3. Musical stimuli 3.1 Listener 3.1.1 Performed by others 3.1.1.1 All elements are predetermined 3.2.2 Emotional experiences – b) Predetermined	Ponto Vista: musical stimuli	Prescribed music
3.Musical stimuli 3.1 Listener 3.1.2 Self-performed 3.1.2.1 Choice of piece – a) Open A) Composed – c) Self; d) Others	Ponto Vista: musical stimuli	Listener is also performer

2.5.1 Overall design: mixed method

A two-group, mixed-method design was used, with both groups being experimental. Since this was an exploratory methodological study, it seemed best to use both qualitative and quantitative methods to obtain and analyse data. Because the participants' profiles differ, and since this study only involved a small group of formally and informally trained musicians, the research could not depend upon statistically significant results. Furthermore, answering the research questions necessitated qualitative research methods, seeing that participants were allowed to provide free descriptions to their own experiences, and individual semi-structured interviews were conducted in order to collect data that could not be gathered through quantitative methods. By using both qualitative and quantitative methods the data could be studied from various perspectives.

It is a methodological study since the main aim is to improve the methodology used in a previous study (Steinberg, 2006). It can be described as exploratory since the two groups tested in this study only consisted of five students each. In the first place, this is a small number of participants, and in the second place the participants themselves cannot necessarily be considered to be representative of the group to which they belong.

2.5.2 Ponto Vista: three Test Periods vs. one Test Period

When comparing listening experience, which is a temporal act, with something like musical score analysis, which is a spatial act, the most noticeable difference is that in listening experiences the musical work is never wholly present. The listener only hears a succession of notes which exists for a moment in time. In contrast, when analysing a score, the musical work can at once be viewed in its entirety; relationships between temporally distant parts of the music can be established. Studies have proven that untrained listeners find it difficult to recognise themes, entry of new musical ideas, or even to identify variations of themes when listening to music for the first time (Spackman and Tan, 2005:135). Listeners often rely on extra-thematic cues like dynamics, texture, articulation and pauses in judgments of music. Recognition of

themes after only one exposure is mostly poor for both trained and untrained listeners; they often need repeated listening experiences or special instructions. There are studies that suggest that even trained listeners initially rely on extra-thematic characteristics, and only focus on thematic and structural features during repeated listening experiences (Spackman and Tan, 2005:134-136). Keeping this in mind, in order for a researcher to capture a truthful representation of an individual's listening experience, irrespective of the strategy they used, it seems clear that repeated exposure to musical stimuli is needed. Although continuous measurement of emotional experiences is not entirely a new concept (Gabrielsson and Lindström, 2001:226-227; Schubert, 2001) it seems as if longitudinal studies using continuous measurement of listening experiences to the same musical stimuli have not been conducted yet. Thus, for this study, three test periods instead of just one was used to address this issue.

2.5.3 Ponto Vista: continuous measurement

The term 'continuous' is used to indicate that responses are measured without interruption. The term has been used consistently in literature (Schubert, 2001:393-394).

Many styles of music express or evoke different emotions as time unfolds. Continuous response tools measure self-reported emotional responses *during* the listening process, instead of afterwards, which help researchers to better understand the variations in responses and the dynamics of listening experiences. In this way a more realistic presentation of listening experiences is possible. This approach contributes to the credibility of the experimental design (Schubert, 2001:393). Continuous measurement enables responses to be measured to an entire piece of music, as opposed to only a sample or movement. Therefore, the musical context can be placed for the listener, and responses can be made in a more realistic listening environment (Schubert, 2001:406).

It is for this reason that continuous measurement of self-reported response to emotional experience of music was used in the current study. It is included in the

methodology as an improvement on the methodology used in previous studies (Steinberg, 2006; Steinberg and Taljaard, 2007).

Continuous measures of self-reported emotional response in music have been used for four broad purposes: validation, comparative investigations, stimulus-response investigations, and system-dynamics investigations (Schubert, 2001:397). In relation to this study, it enabled the researcher to conduct a comparative investigation between formally and informally trained musicians. However, the most important purpose is a better understanding of the time-dependent, dynamic nature of the music-emotion system (Schubert, 2001:399). One of the benefits of continuous measures of emotion in music is to be able to determine the *reason* for changes in responses as a function of the music (Schubert, 2001:404). Although this is not the main aim of this study, it could be used in the future in relation to participants' listening profiles, to determine the reason for a participant's responses.

Some of the most common practices of continuous measurement have also been used as basis in this study. The most common method today of synchronizing participants' responses and the music they listen to with each other is done by a computer-generated clock which the software uses to track the location of the music and is used as a tag for the participants' responses (Schubert, 2001:403). The questionnaire used in Part II of the methodology was therefore computerised into a program called 'Ponto Vista' in order for the researcher to be able to monitor the response rate, time and manner of the participants. It also enabled the researcher to synchronise any selections the participants made from the various checklists in Question 2 with the music they listened to. One of the response types also incorporated in this study is an inertia response called *after-glow*. After-glow is where the participant still experiences an emotion in the music, even though the music has finished (Schubert, 2001:404). It is from this response type that the idea of the fifteen-second window was derived. At the end of each musical track, the program allows the participant an extra fifteen seconds during which he/she can make any last selections from the various checklists. More information on the program is provided in Chapter 4.

2.5.4 Ponto Vista: self-report

Self-report response formats, and not observer reports (except for the Tomatis listening test), are used in this study since that is the common practice in measuring emotional experience of music (Gabrielsson and Lindström, 2001:224-235; Schubert, 2001:394).

2.5.5 Ponto Vista: verbal and non-verbal response formats

As mentioned before, both verbal and non-verbal testing formats are included in this study. Non-verbal testing formats are included as an improvement of the methodology. Not all people are equally skilled at communicating their emotions in words. They are capable of experiencing them, but not always able to communicate those emotions in words, let alone do so in an understandable way. For this reason non-verbal response formats were included in this study in order to present a participant with more options/ways in which to communicate emotional experiences that suits his/her personality and capability. The concept of using non-verbal response formats to gather data has been used in other studies as well (Gabrielsson and Lindström, 2001:227).

The verbal testing formats used in this study consist of word sorting, a word checklist, free description, and rating scales of psychophysical parameters. Checklists consisting of words, descriptive terms or nouns, as well as free description and rating scales are three traditional self-report formats used in the measurement of the emotional experience of music (Gabrielsson and Lindström, 2001:224). Checklists and rating scales are also used in continuous measurement (Schubert, 2001:394). Word sorting was added to the methodology to present the participants with an opportunity to construct their own word categories. This is also consistent with the procedure that Shaver and colleagues followed in their study (Shaver *et al.*, 1987:1065). The three traditional formats (word checklist, free description and rating scales) together with word sorting constitute the four questions in the Ponto Vista program.

Non-verbal testing formats used in this study consist of a colours checklist and facial expressions checklist. The relationship between music and colour has been studied

before (e.g. Davies, 1980:86; Viora, 1982:335-336). A study on the relationship between facial expression, colour and basic emotion is also worthy of mention (Da Pos and Green-Armytage, 2007). Facial expressions have also been used in studies investigating the emotional experience of music (e.g. Gabrielsson and Lindström, 2001:234). The relationship between emotion and facial expression has also been investigated (e.g. Ekman, 1992; Ruch, 1995; Constantini *et al.*, 2005). Words and pictures of faces expressing emotions have been used in continuous measurement studies (e.g. Schubert, 2001:402). The facial expressions and colours used in this study were set in checklists as part of Question 2 (together with the word checklist) in order for participants to use all these formats during the question to indicate their emotional experience of the music.

2.5.6 Musical stimuli: prescribed music

Sloboda and Juslin (2001:93) state that several views on the expressive capabilities of music have come into existence. Amongst others, music has been viewed to express objects, persons, events, political and even social events or conditions. This might have contributed to the theory of universalism as described by Davies (2001:37) which created the concept that music is a 'universal language of emotion' and that musical and emotional expressiveness can be recognised cross-culturally.

The raga, the most central aspect of Hindustani classical music, can be described in terms of a specific succession of scale degrees from which it is derived as well its ascending and descending patterns, which are fixed. It is subjected to rules that govern the melodic movements and is limited to a specific collection of notes and ways in which these notes can be ornamented or emphasized. The word 'raga' means to be coloured, affected, moved, charmed, or delighted, and is that quality of an object that renders emotion. This aesthetic aspect determines the uniqueness of every raga through the particular notes and melodic movements within the structure of a raga. The performer of a raga is expected to express an aesthetic 'flavour' during the performance. This is usually a fixed, predetermined emotion, and can be evoked in a sensitive listener through his 'tasting' of the symbolised emotion in the raga (Dick *et al.*, 2001:220; Powers and Widdess, 2001:178-179; Magriel, 1998).

One can test whether emotional expressiveness can indeed be perceived cross-culturally by exposing listeners to Hindustani ragas without revealing the predetermined emotion to them. Since Hindustani music is associated with a given set of expectations, its employment in this study seems appropriate.

2.5.7 Musical stimuli: listener is also performer

Persson (2001:275) states that music as a cultural phenomenon would most likely lose its universal appeal if its potential to arouse some kind of emotional response is misplaced. He further suggests that emotions, feelings and affects are mostly communicated in the performance of and listening to music, and that these issues are vital when attempting to understand music.

The ideals of music education in the training of performers to express emotion in musical performances as expressed by Juntunen (2004) and Swanwick (1999) have already been referred to in Chapter 1. Most studies connected with the emotional experience of music use the following as musical stimuli:

1) selected pieces of music
2) entire pieces of music
3) short sound sequences without musical context subjected to systematical variation of structural factors
4) real music subjected to variation of structural factors
5) phrases from known pieces
6) newly composed musical phrases
7) a variety of intervals, chords or patterns
8) synthesised versions of known phrases or pieces
9) rearranging the parts of a composition, the order of movements or variations of known musical works
10) popular, folk or traditional music

The musical stimuli is usually in the form of an existing recording as performed by professional musicians, but live performances and synthesised computer sampling

have also been used. As far as can be determined, the music is not performed by the participants themselves (Gabrielsson and Lindström, 2001:224-234).

It seems as if the communication of emotion in music can only exist between a performer and a listener through the vehicle of music performance. A performer must learn to effectively communicate emotion in a performance in order for a listener to experience the music as appealing. It seems then as if the focus should be more on the performer and the development of performing abilities. The current study specifically focused on performing musicians (or students who perform on a regular basis) in order to determine, amongst other things, whether or not music education has been effective in teaching them how to communicate emotion to an audience during a performance. The focus was also to determine whether or not performing musicians are able to experience emotion themselves, since as C.P.E. Bach argued "a musician cannot move others unless he too is moved" (Persson, 2001:275). There are also few studies that focus on the emotional responses of music performed *and* perceived by the same person. As Persson (2001:285) has put it "the phenomenological excursion into the subjectivity of musical performers has generated a number of hypotheses that need testing and further exploration in order to arrive at verifiable cognitive models…".

As has already been mentioned, informally trained musicians seem to have the opportunity to emotionally experience their music much more. They have more freedom to experiment, play together in bands, and perform the music they like and can identify with (Green, 2002:104-107). It seems as if informally trained musicians have more opportunities than formally trained ones to enjoy their own performances (Green, 2002:99-107). According to statistics, informally trained musicians are in the majority (Green, 2002:1, 2, 5). Since Green's study was done in Britain, the statistics only reflect on British adults, but one can safely assume for the purposes of this current study that the same is true for South African musicians.

For the current study the participants were allowed to select any two pieces to perform themselves. Some participants chose to perform their own compositions. Other participants decided to select pieces from the standard B. Mus-repertoire.

Since the current study focused on own performances and the performer as listener (amongst other things) it seemed important to compare the emotional experiences of formally and informally trained musicians. Music educators could possibly learn from informally trained musicians and assess their teaching methods in order to obtain the goal of teaching: the ever better performance.

2.6 SUMMARY

This chapter contained descriptions of methods of empirical research conducted in relation to music and emotion. Existing methods as well as potential methods that could be used in future research were presented in this chapter. It furthermore contained a description of the studies and measuring instruments upon which the research design of the study described in this research report were based, as well as a discussion on other method-related decisions. This chapter therefore answered the first two sub-questions of this study.

The next chapter presents information on the empirical study regarding the research design, methodology and measuring instruments, participants, recordings, procedure, data capturing and editing.