Improving Specificity of the Musical Mood Induction Procedure

By Lisa Liskovoi

Hon B.A., Bishop's University, 2009

A thesis presented to Ryerson University in partial fulfillment of the requirements for the degree of Master of Arts

in the program of Psychology

Toronto, Ontario, Canada 2013

Author's Declaration

I hereby declare that I am the sole author of this research paper. This is a true copy of the research paper, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this research paper to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this research paper by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my research paper may be made electronically available to the public.

Improving Specificity of the Musical Mood Induction Procedure

By Lisa Liskovoi

Ryerson University, Master of Arts, Psychology, 2014

Abstract

The musical mood induction procedure was used to induce 3 negative moods: sadness, fatigue and anxiety. Induction was validated using subjective and physiological measures. One hundred twenty-seven participants listened to one of 18 film soundtrack excerpts for 20 minutes. Physiological response (heart rate, respiration, skin conductance level (SCL), and facial electromyography) was recorded throughout the induction and postinduction phases. Subjective mood ratings (sadness, anxiety, tiredness, valence, arousal) were provided before induction and throughout the postinduction phase. Repeated measures ANOVAs showed increase in valence and decrease in arousal in all conditions after induction, which persisted in the postinduction phase, and an increase in tiredness immediately after induction. Reduction in SCL was strongest in the fatigue condition. However, difference between groups was only evident when comparing fatigue and sadness conditions between 3-10 minutes. Lack of between-group differences and mixed physiological findings suggest that specificity is difficult to achieve through musical mood induction.

Acknowledgements

I would like to thank Dr. Frank Russo for being an incredibly supportive and patient advisor, and for providing helpful guidance and advice. I would also like to thank Chris Lachine for his work compiling the music for the project, Salma Shaikh for her assistance in data collection, Dr. Naresh Vempala for his input and assistance, and Gabe Nespoli for his help with data processing. I'm thankful to my family, friends and lab-mates for their encouragement and support. Lastly, I would like to thank the members of my thesis supervisory committee; Drs. Colleen Carney and Naomi Koerner, as well as Dr. Julia Spaniol, my external examiner, for their thoughtful comments and suggestions.

Table of Contents

List of Tables

List of Figures

List of Appendices

Improving Specificity of the Musical Mood Induction Procedure

Mood plays a central role in every part of our lives. How we feel can govern the way we think, behave and perceive the outside world. As a result, it is vital for psychologists to understand the effects of different moods on our psyche. In order to gain scientific understanding of such states, researchers must first be able to reliably manipulate it in a controlled setting.

Mood induction procedures are methodologically important in many research areas, including social, personality and clinical psychology (Västfjäll, 2002). A variety of methods have been used to date, such as Velten (a technique that presents participants with a series of self-referential statements), imagination, and film/story. However, *musical* mood induction offers some advantages over alternative methods. The "[moods] instill[ed by musical mood induction] are reasonably strong, stable, sincere, and reproducible" (Eich et al, 2007, p. 133). The effectiveness of this procedure is evident in self-report, behavioural and physiological changes (Västfjäll, 2002; Clark, 1983; Eich et al, 2007, Bernardi, Porta & Sleight, 2006). However, limitations exist in the utilization of this procedure. In particular, the musical excerpts chosen for induction are often selected on intuitive grounds without prior validation (Eich et al, 2007), and the specificity of moods that have been successfully elicited is narrow (Martin, 1990).

The present project was designed to address these limitations. In order to improve the specificity of mood induction, I chose musical excerpts to induce three moods that are broadly classified as negative but distinct with regard to the valence and arousal dimensions; fatigue, sadness, and anxiety. I then used physiological and subjective response to these excerpts for the purpose of validation. Based on findings that individuals' responses are

mediated by their ability and willingness to be absorbed in the music, I chose to exclude participants low on this trait from the analysis. I hypothesized that the three mood conditions would elicit unique subjective and physiological profiles.

Defining affective terms

In the context of music, affect can be conceptualized along a time-dependent continuum, ranging from perception to felt emotion to mood. Perceived emotions refer to the affective content the music is believed to represent, which can be accurately identified within milliseconds (Peretz, Gangnon & Bouchard, 1998). Felt emotions are brief affective responses, typically evoked by a few seconds or minutes of music. Moods are induced by longer exposure to music, around 10 minutes or more (Mongrain and Trambakoulos, 2007; Pigantiello, Camp & Rasar, 1986), are more prolonged, and include a wider range of affective states.

Listeners can accurately perceive affective content in musical excerpts as short as 250 milliseconds (Peretz, Gangnon & Bouchard, 1998). When musical presentation is short, individuals can identify the emotional intention of music, but may not experience an emotional reaction. When music is presented for longer segments of time, such as several seconds to a few minutes, listeners are likely to experience felt emotions. Emotions are brief, intense, and object-directed affective episodes (Grey and Watson, 2007).

Moods are prolonged states that include a wide range of affective experience. Whereas emotions last seconds or minutes, moods can last several hours or even days. Emotion definitions emphasize intense, high-activation states. As a result, traditional emotion models have typically excluded low-activation states such as serenity. In contrast, moods refer to all feeling states, including states that are not normally regarded as emotions, such as

fatigue (Watson, 2000). Notably, fatigue is an important part of the human affective experience, and is included as a core dimension in the Profile of Mood States (POMS)–one of the most widely used mood scales in psychological literature (Lieberman, Bathalon, Falco, Kramer, Morgan & Niro, 2005).

Although a relationship between perception of intended affective content and experience of affective response exists, music that conveys a particular emotion does not necessarily elicit it. Direct comparisons between the perceived emotional content of music and the felt emotion caused by it have shown varying degrees of agreement. Zentner and Scherer (2008) found a significant difference between the reported frequency of felt and perceived emotions in response to music of different genres. Evans and Schubert (2008) reported that perceived and felt ratings of pleasantness and arousal were congruent in 61% of cases. Kallinen and Ravaja (2006) reported that even though perceived and felt ratings of pleasantness and arousal in music tended to be in the same direction, the magnitude of ratings was significantly stronger for perceived than for felt emotions. These findings suggest that the apparent emotional content of a particular musical excerpt is not necessarily a valid predictor of the emotion it will elicit in the listener. Furthermore, to the best of our knowledge, no studies have compared the ability of perceived emotion or brief felt emotion to evolve into a more stable mood state with longer exposure. Thus, the effectiveness of excerpts at eliciting specific moods needs to be validated empirically.

Assessment

The present study evaluated musical excerpts used for mood induction both subjectively and physiologically. Both dimensional and discrete self-report measures were used in order to quantify subjective affective response. Heart rate, skin conductance level

(SCL), and respiration, collectively gauging autonomic system activation, were used to assess physiological changes in arousal. Facial electromyography was measured in order to capture the physiological experience of valence.

Several models have been used to assess subjective affective response, which can be broadly grouped as either discrete or dimensional. Discrete measures ask participants to describe musical excerpts using specific emotion labels such as happiness or anger. Dimensional scales, on the other hand, require subjects to define their moods in terms of dimensions such as valence (positive or negative) and arousal (excited or calm). In a comparison of these two approaches in the context of music, Vuoskoski and Eerola (2011) found that the dimensional model was superior in inter-rater agreement and accounted for 89.9% of the variance in ratings on all scales.

In the present study, dimensional ratings were collected in order to establish how distinct the three negative moods are in the valence/arousal space. Since the study aimed to induce highly specific negative moods which may exhibit some overlap on these dimensions, particularly valence, continuous ratings of tiredness, sadness and anxiety were collected as well.

Measures of peripheral physiology provide a noninvasive method for assessing bodily reactions to mood induction and changes in affect. SCL is a measure of sympathetic nervous system activation (Caccioppo, Tassinary & Bernston, 2007), and has been shown to demonstrate high effect sizes in differentiation of neutral from anxious states in experimental settings (Blechert, Lajtman, Michael, Margraf & Wilhem, 2006). Since the function of the sympathetic nervous system is also central in alertness and has been found to increase with arousal (Zimny & Weidenfeller, 1963), this measure is appropriate for gauging the

effectiveness of both fatigued and anxious mood state induction. Heart rate has also been linked with arousal, and thus can be used to differentiate high and low arousal states such as fatigue and anxiety (VanOyen Witvliet & Vrana, 2007, Lang, Greenwald, Bradley & Hamm, 1993).

Facial electromyography can be used as a physiological indicator of experienced valence. Increased corrugator activity has been linked to stimuli that are negative in valence (VanOyen Witvliet & Vrana, 2007, Lang, Greenwald, Bradley & Hamm, 1993) and low in arousal (VanOyen Witvliet & Vrana, 2007). In contrast, zygomatic activity has been linked to positively valenced (pleasant) and highly arousing stimuli. Both corrugator and zygomatic activity are more strongly linked to valence than arousal. (VanOyen Witvliet & Vrana, 2007, Lang, Greenwald, Bradley & Hamm, 1993).

Review of MMIP validation

Studies have shown that musical mood induction causes changes in self-report, behaviour and physiology (Västfjäll, 2002; Clark, 1983; Eich et al, 2007, Bernardi, Porta & Sleight, 2006; Lundqvist, Carlsson, Hilmersson & Juslin, 2009). Several reviews have documented the high success rates for musical mood induction compared to other procedures. In one study comparing Velten and musical mood induction methods, 87% of participants reported that they experienced a genuine change in felt emotion after musical mood induction, compared to 50% in the Velten condition (Clark and Teasdale, 1982, as cited in Clark, 1983.) Hypnotic suggestion methods have reported success rates as low as 15% (Eich, 2007), while success rates for MMIPs have been reported as high as 75% (Martin, 1990) and 100% (Gerrards-Hesse, Spies & Hesse, 1994; Clark, 1983).

However, Westermann, Stahl, and Hesse (1996) point out that definitions of

successful induction vary between studies, and in their meta-analysis of mood induction procedures, MMIPs had a combined effect size of merely .32.

One possible reason for such variability in success rates may be the difference in methodology throughout studies. Some studies have used within–subject comparisons, others have compared changes in mood between conditions. Not all studies have measured mood at baseline, and among those that have, there is no consensus regarding the way it is used in data analysis. A closer look at the methodology of several studies that have assessed the effectiveness of musical mood induction can help illustrate the similarities and differences of induction procedures. Only studies that sought to specifically investigate the effects of music on changes in mood, related behaviours and physiological responses will be discussed in detail.

In a study by Mongrain and Trambakoulos (2007) the researchers used a repeated measures design, first inducing sadness and then happiness. Participants provided mood ratings–using VAS scales of dysphoria and positive affect–four times; at baseline, after 10 minutes of sad music, after completing questionnaires in between the two types of induction, and after listening to 10 minutes of happy music. Results showed significant changes in mood after induction. Ratings of dysphoria increased by 36% while ratings of positive affect decreased by 19% compared to baseline. Similarly, ratings of dysphoria decreased by 41% while positive affect increased by 23% after positive mood induction.

In addition to self-report, several studies have included behavioural tasks to validate mood induction. Following induction of negative mood, participants tend to exhibit an array of behaviours that are characteristic of depression such as psychomotor retardation (Teasdale, 1982 and Richards, 1981, as cited in Clark, 1983) and loss of incentive (Clark and Teasdale,

as cited in Clark, 1983). Kenealy (1988) assessed differences in mood between happy, sad and neutral conditions using the Multiple Affect Adjective Checklist (MAACL), six line scales (0-100) pertaining to different mood adjectives, and several behavioural tasks (writing speed, decision time, word association, distance approximation). After listening to music for eight minutes, scores were significantly different on four of the six line scales (happy, exhilarated, sad and despondent), as well as the MAACL depression scale, between happy and sad conditions. Between-group comparisons revealed significant differences between the happy and sad conditions in the expected direction on all behavioural tasks except word association. Participants in the happy condition provided higher distance approximation values, had faster writing speed, and quicker decision times.

A study by Pignatiello, Camp and Rasar (1986) compared participants in elated, depressed, and neutral conditions after 20 minutes of music listening using the Depresssion Adjective Checklist (DACL) and a psychomotor task (writing backwards from 100). Compilation tapes of popular, soundtrack, and classical music were created for each of the three moods on the basis of characteristics such as rhythm, mode, loudness, melody and tempo. Participants were asked to listen carefully, but were not informed that the music was intended to alter their moods. Those in the depressed group had significantly higher DACL scores, and performed significantly poorer on the psychomotor task, consistent with hypotheses. Although these two studies were able to show differences between conditions, participants did not provide baseline mood ratings in either study, making it difficult to assess the change in mood caused by the music.

To address the lack of baseline in their first study, Pignatiello, Camp and Rasar (1986) conducted a follow-up study identical to the first one, but included a pre-induction

psychomotor task. Scores on this task were then used as a covariate in the analysis of post-induction results. The elated group performed significantly better on the psychomotor task compared to the depressed group, even when baseline scores were taken into consideration. Unfortunately, they did not measure baseline levels on the DACL.

Fewer examples of musical mood induction of specific moods exist in the literature. In addition to happy and sad moods, Mayer, Allen, and Beauregard (1995) attempted to elicit anger and fear through music and vignettes and validate the procedure in three studies. In the first study, ten music judges rated 20 nonvocal classical music pieces on five-point scales for each of the target moods. Ratings were then averaged across judges, and pieces that were relatively high on target mood and low on nontarget moods (no specific criteria given) were selected. In the second study, 36 participants experienced all mood inductions in a between-subjects design, first listening to one minute of music and then reading eight corresponding vignettes for each mood. Participants were asked to complete a 16-item mood adjective scale containing four synonyms of the target moods, once at baseline and then after each induction. Inductions were successful at enhancing target moods from baseline, and further analysis of negative moods revealed that inductions raised target moods while causing only slight increases in nontarget moods. The final study sought to validate the induction of happy and sad moods using a judgment task. Anger and fear were dropped from the last experiment. After undergoing musical mood induction, participants completed two pleasant-unpleasant judgment scales. Findings were congruent with predictions, with happy induction causing more positive judgments than sad induction.

A study by Albersnagel (1987) attempted to elicit anxious, depressive, elated and neutral moods. He exposed participants to 20 minutes of music, selected by a musicologist to

elicit one of the target moods. Induction measures included VAS scales of discrete target moods, and a word association task. Although the study failed to show increases in anxiety-related thought associations, significant increases in anxiety ratings were evident after mood induction.

Physiological changes following music listening provide further support for the validity of MMIPs. Using a method identical to the one in the aforementioned Pignatiello et al 1986 study, the researchers investigated the influence of depressed and elated music on blood pressure, heart rate, finger pulse amplitude and respiration rate. Results showed that participants in the elated condition experienced increased heart rate, and systolic blood pressure (Pignatiello, Camp, Elder and Rasar 1989, as cited in Västfjäll, 2002).

Several additional studies have assessed the effects of brief musical excerpts on emotional response to music using physiological measures. Lundqvist, Carlsson, Hilmersson & Juslin (2009) presented listeners with happy and sad pop songs composed specifically for the study, ranging from 140 to 204 seconds, while measuring facial EMG activity, heart rate, skin conductance, and finger temperature. Positive moods lead to increased zygomatic facial muscle activity and greater SCL. Another study measured a range of cardiovascular, cerebrovascular and respiratory indices while participants listened to two-minute excerpts of music of different genres. Music with faster tempi and simple rhythmic structures tended to increase ventilation, blood pressure, and heart rate, while simultaneously decreasing mid-cerebral artery flow velocity and baroflex sensitivity (Bernardi, Porta and Sleight, 2012). Sandstrom and Russo (2010) found that music rated high on valence and low on arousal was effective at decreasing SCL and heart rate following an acute stressor. These findings suggest that music is not only able to alter perceived mood, it also alters a listener's physiological

state in a manner that is theoretically consistent with the induced emotion.

Despite the existence of validation studies, an extensive review of music and emotion studies done by Eerola and Vuoskoski (2013) revealed that merely 9% of researchers chose musical excerpts based on previous research. Of the remaining studies 33% reported no information regarding the selection method, 39% chose researcher-selected music, 8% based selection on a pilot study, 6% selected music using a panel of experts, and 4% allowed participants to choose their own music. Although the review covered a range of music and affect literature, and did not focus on musical mood induction, it is evident that the vast majority of music is not validated prior to use. In addition, to the best of our knowledge, there have been no successful attempts to induce fatigue through music, and the procedure calls for improvements in specificity.

Classification

Selecting appropriate moods for the purpose of improving specificity of musical mood induction requires a theoretical understanding of affect classification. Two prevailing approaches for understanding moods have been used extensively in music and affect literature: discrete and dimensional (Juslin & Sloboda, 2010). Discrete models, such as the basic emotion model (Ekman, 1992; Ekman, 1999), group emotions into distinct categories (i.e. happy, sad, fearful, angry). In contrast, dimensional models organize affect along continuums within an n-dimensional space (Russo, Vempala & Sandstrom, 2013). Perhaps the most noted dimensional model of affect is Russell's circumplex model (Russell, 1980). The circumplex model of affect organizes moods along a vertical dimension of arousal, ranging from calm to excited, and a horizontal dimension of valence, ranging from unpleasant to pleasant. Refer to Figure 1 for a visual representation of Russell's circumplex

model, and the location of target moods in the present study within the valence/arousal space.

The three moods selected for the study occupy the broad category of negative affect but exhibit subtle differences on dimensions of valence and arousal, captured in Russell's circumplex model. Numerous studies have classified sadness, fatigue, and anxiety as moods within the valence/arousal model (Vuoskosi & Eerola, 2011; Sandstrom & Russo, 2010; Feldman Barrett & Russell, 1998; Kawahara & Sato, 2013; Russell, 2003; Posner, Russell & Peterson, 2005).

Sadness

Based on Russell's model (Russell, 1980), sadness is low in valence, higher in arousal than fatigue and lower in arousal than states such as tenseness and distress which are similar to anxiety (Vuoskosi & Eerola, 2011; Sandstrom & Russo, 2010). Sadness causes individuals to judge people and situations more negatively (Wehmer and Izard, 1962, Izard, 1964, Wessman and Ricks, 1967, as cited in Forgas, 1995), and increases recall for negatively valenced content (Bower & Forgas, 2001). Following musical induction of sad mood, participants have shown bias towards recognition of negatively valenced words and impaired facial emotional recognition, compared to a neutral condition (Chepenik, Farah, & Cornew, 2007).

Mixed results have been found for autonomic nervous system response for sadness (Kreibig, 2010). Depending on the type of induction, sadness may cause an "activating" or a "deactivating" response. An activating response is generally characterized by increased heart rate, increased or unchanged SCL, and variable respiration rate changes, with some studies reporting increases while others reporting decreases. A deactivating response, on the other hand, is consistently characterized by decreases in heart rate, while showing more mixed

results for respiration rate and SCL, reporting increases in some studies and decreases in others.

According to Kreibig (2010), activating response to sadness has been reported in mood induction studies that have used directed facial action, personalized recall and some film materials, while the deactivating response has been reported in most studies that have used film material, studies that utilized standardized imagery, and those that used musical excerpts. Based on Kreibig's analysis, it is more likely that the sadness elicited in the present study would be deactivating in nature.

Fatigue

Within the affective space of Russell's circumplex, fatigue is lowest in arousal and has neutral or slightly negative valence (Feldman Barrett & Russell, 1998; Kawahara & Sato, 2013). Fatigue can manifest both physiologically and cognitively (Sharpe & Wilks, 2002; Carney, Moss, Lachowski & Atwood, 2010). Physiologicaly, fatigue is experienced as a lack of energy (Sharpe, Chalder, Palmer & Wessely, 1997). It has been linked to autonomic system dysfunction in some conditions, such as Multiple Sclerosis and Chronic Fatigue Syndrome (Niepel, Bibani, Vallisaar, Langley & Bradshaw, 2013; Pagani & Lucini, 1999). However, the relationship between autonomic system activity and fatigue is complicated. In Chronic Fatigue Syndrome, fatigue symptoms have been linked to hypoactivity of the sympathetic (activating) component of the autonomic nervous system (Pagani, & Lucini, 1999). In patients with Multiple Sclerosis, on the other hand, fatigue has been linked to reduced sympathetic responsivity (Niepel, Bibani, Vallisaar, Langley & Bradshaw, 2013). Healthy adults who have reported experiencing symptoms of fatigue after exercise deprivation have exhibited baseline hypoactivity of the autonomic nervous system. These

mixed findings regarding the relationship between autonomic system activation and fatigue are further complicated by it being unclear how activation is manifested during actual experience of fatigue.

Cognitive symptoms of fatigue include inability to cope with sustained mental tasks, difficulty concentrating, and reduced alertness (Chaudhuri & Behan, 2000; Carney, Moss, Lachowski & Atwood, 2010). Fatigue (in the absence of insufficient sleep) has been linked to heightened levels of anxiety, worry and feelings of insecurity (Montgomery, 1983). However, the direction of the relationship between fatigue and these negative cognitions has not been established, and evidence for the experimental induction of fatigue is lacking in the literature. This is partly because it is unclear whether it is possible to systematically elicit fatigue in the lab. It is important to note that fatigue is distinct from sleepiness and therefore cannot be induced via sleep deprivation since it often occurs in individuals who receive an adequate amount of sleep (Shen, Barbera & Shapiro, 2006; Montgomery, 1983). As a result, musical mood induction makes a suitable candidate for attempting to experimentally induce fatigue.

Anxiety

Anxiety is high in arousal, and low in valence in Russell's circumplex of affect (Russell, 2003; Posner, Russell & Peterson, 2005). Anxiety has been found to affect cognitive processes in a number of ways. Anxious states are associated with attentional biases towards threatening stimuli (Bar-Haim et al, 2007). Individuals experiencing anxiety report more task-irrelevant thoughts, in social situations as well as during tests, and exhibit relatively poor test performance (Sarason, 1990). Induced anxiety has also been found to

impair both verbal and spatial working memory (Vytal, Cornwell, Letkiewicz, Arkin & Grillon, 2013).

A review of research that has monitored automonic nervous system responses during different affective states found that induction of anxious mood consistently increases sympathetic nervous system activation, resulting in increased heart rate, faster respiration rate and higher SCL (Kreibig, 2010).

Absorption in Music

Research suggests that absorption—the capacity and willingness to be drawn in by a stimulus (Sandstrom & Russo, 2013)—plays a significant role in affective response to music. In a study by Sandstrom and Russo (2010), absorption as measured by the Tellegen Absorption Scale (Tellegen & Atkinson 1974) was found to be a significant predictor of physiological recovery from stress. This finding caused Sandstrom and Russo (2011) to develop the Absorption in Music Scale (AIMS), a scale that adapts the construct of absorption to the realm of music listening. They found that high absorbers showed stronger difference scores in valence, indicating a more pronounced change in mood after listening to music. Based on these findings, participants in the current study were asked to complete the AIMS scale so that those low in music absorption could be eliminated from the analysis because they were expected to exhibit lower reactivity to music. For the purposes of subjective and physiological measures, data were dropped from participants scoring in the lowest quartile of absorbers (AIMS< 97).

Hypotheses

Based on the predicted locations of sadness, tiredness, and anxiety within Russell's circumplex, I hypothesized that the three states would elicit unique subjective and

physiological response profiles. Seeing as the induction was of mood states, which are prolonged in nature, rather than an emotional response, observed changes in mood caused by the music were predicted to remain stable after induction.

I predicted that participants would report a decrease in pleasantness post induction irrespective of the assigned condition and that the pattern of pleasantness and arousal ratings would vary across conditions. Specifically, I hypothesized that excerpts in the fatigue condition would yield the strongest decrease in arousal compared to baseline, and highest pleasantness and tiredness ratings compared to the sadness and anxiety conditions. Physiologically, I hypothesized that fatigue would manifest in lowest heart rate, respiration, and SCL, as well as lowest corrugator activity and highest zygomatic activity.

I predicted that the anxiety condition would exhibit the highest increase in arousal ratings, smaller increase in pleasantness ratings than the sad condition, and larger increase in pleasantness ratings than the fatigue condition. I also predicted that physiological response to anxious music would result in increases in heart rate, respiration and skin conductance, increased corrugator and decreased zygomatic activity over time.

I predicted that the sad condition would report greatest reduction in valence post induction, smaller decrease in arousal levels than the fatigue, and a decrease in arousal compared to the anxiety condition. It was more challenging to make predictions regarding physiological response based on mixed findings in the literature and the required distinction between activating and deactivating sadness. However, based on Kreibig's (2010) findings that music tends to cause a deactivating sadness response, I predicted that listening to sad music would result in decreased heart rate and SCL. Based on the aforementioned physiological findings reporting a link between measures of valence and activation of the

corrugator and zygomatic muscles, I further predicted that sad music would result in increased corrugator, and decreased zygomatic activity.

Method

Participants

One hundred and twenty-seven participants were recruited through the Ryerson University SONA undergraduate student participant pool and were awarded one course credit for their participation. Refer to Table 1 for a full summary of demographic statistics for the sample. One hundred and seventeen participants were right handed, eight were left handed, and two were ambidextrous but used their right hand for writing. Forty-one participants were randomly assigned to the sadness condition, 46 to the fatigue condition, and 40 to the anxiety condition. Refer to Table 2 for number of participants per each track.

Apparatus and Materials

All participants were given a consent form, demographic information sheet (e.g., age and gender), the Depression, Anxiety and Stress Scale (DASS21), the Epworth Sleepiness Scale, the Fatigue Severity Scale (FAS) and the Absorption in Music Scale (AIMS). Since the study is addressing the induction of sadness, anxiety and fatigue, these scales were meant to ensure that changes in mood cannot be explained by variation in these affective states in their everyday lives. Finally, participants received a booklet of questions assessing subjective mood response, and a questionnaire evaluating liking, familiarity, and perceived predominance in the musical track.

Measures

The Depression, Anxiety and Stress Scale (DASS21) is a 21-item scale that assesses depression, anxiety and stress (Lovibond, & Lovibond, 1995). Henry and Crawford (2005) reported Cronbach's alpha values of .88 for the depression scale, .82 for the anxiety scale, .90 for the stress scale and .93 for the total scale, indicating good reliability. Convergent validity for the full-length version of the scale has been assessed using Pearson product-moment correlations of the DASS and the Hospital Anxiety and Depression Scale (HADS) as well as the Personal Disturbance scale (sAD). The DASS depression scale correlated with the sAD depression scale, $r = .78$, and with the HADS depression scale, $r = .66$. The DASS anxiety scale correlated with the sAD anxiety scale, $r = .72$, and with the HADS anxiety scale, $r = .67$ (Henry & Crawford, 2003). In a later study, the authors reported comparable convergent validity for the abbreviated version of the scale (Henry & Crawford, 2005). The Anxiety and Depression subscales were used to assess the contribution of pre-existing anxiety, depression, and sadness to mood change.

The Fatigue Severity Scale (FSS) (Krupp et al, 1989) is a 9-item scale that assesses fatigue level. The measure has high internal consistency, with Cronbach's alpha of .88 for participants healthy participants (Krupp et al, 1989 as cited in Lichstein et al, 1997), as well as good test-retest reliability, $r = .84$ (Dittner, Wessely & Brown, 2010). The validity of this measure is indicated by higher scores in patients with pathological fatigue (e.g., those with multiple sclerosis or systemic lupus) than normal controls, and significant correlations of FSS scores with scores on a previously established measure of fatigue (Shahid, Wilkinson, Marcu & Shapiro, 2012). The total score of the FSS was used to examine whether changes in tiredness ratings are mediated by preexisting fatigue.

The Epworth Sleepiness Scale (ESS) is a 10-item scale that assesses whether participants are suffering from sleepiness **(**Johns, 1991). Estimated using Cronbach's alpha, the ESS has a reliability coefficient of .88 and good criterion-related validity as evident by a significant correlation with sleep latency in patients with various sleep disorders; r = -.514, p > 0.01 (Johns, 1992). The ESS was included to test whether effects of induced fatigue may be related to experienced sleepiness.

The Absorption in Music Scale (AIMS) is a 34-item scale that measures the extent to which participants are able and willing to become emotionally responsive to music **(**Sandstrom & Russo, 2011). The scale is reliable as indicated by test-retest reliability coefficient of .91. The scale also exhibits good convergent validity, correlating significantly with other scales measuring absorption such as the Tellegen Absorption scale (r = .76, p < .01), which has previously been associated with stronger physiological response to music, and the Musical Involvement Scale (r = .74, p < .01). The AIMS scale was used to account for individual differences between those participants who are more emotionally responsive and those more resistant to absorption in music.

The Visual Analog Scale (VAS) Two types of visual analog scales were used: the Self Assessment Manekin (SAM), and mood specific line scales. The SAM scale is a nonverbal assessment tool of valence and arousal (Bradley and Lang, 1994). The valence and arousal scales are comprised of nine discrete illustrations of different levels of excitement (ranging from *very calm* on the left hand side and *very excited* on the right hand side), and valence (ranging from *sad* on the left hand side, *happy* on the right hand side). The measure's valence scale exhibits high correlations with pleasure factor scores derived from the Semantic Differential ratings system, (r = .96 for paper and pencil, and .97 for computerized

version), and the SAM arousal scale exhibits high correlations with the arousal factor scores of the Semantic Differential ratings system (r = .95 for paper and pencil, and .94 for computerized version) (Bradley and Lang, 1994). In addition, ratings of valence and arousal on the SAM scale show significant correlations with congruent physiological response. Frowning, measured by corrugator muscle activity, is negatively correlated with valence ratings (r = -.9, linear), while smiling, measured by zygomatic muscle activation, and valence ratings are positively correlated (r = .9, quadratic). Skin conductance response is positively correlated with arousal ratings (r = .81, linear) (Lang, Greenwald, Bradley & Hamm, 1993). The scale is also highly reliable as indicated by high split-half coefficients for both valence (rs = .94), and arousal (rs = .94) (Lang, Bradley & Cuthbert, 1997). The other three VAS scales asked participants to mark their level of sadness, tiredness and anxiety along an unmarked line ranging from *not at all* on the leftmost end and *extremely* on the rightmost end. The lines were 100mm long and scored on a scale of 0-100. The absence of numerical values was meant to encourage participants to give a more accurate response from trial to trial rather than relying on previously provided values. Subjects were provided with 22 numbered excerpts of paper containing five VAS scales.

Physiological Measures were recorded with AcqKnowledge software through the Biopac MP150 System using the following BioNomadix receivers and corresponding wireless transmitters: PPGED-R - photoplethysmography (PPG)/electrodermal activity (EDA), EMG2-R - electromyography (EMG), RSPEC-R - respiration (RSP), and UIM100C - analogue audio signal. Sampling rate within AcqKnowledge was set to 1000Hz. NuPrep skin prep gel was used to clean the skin prior to electrode attachment.

One-inch disposable square cloth solid gel EL504 electrodes were used to record

facial EMG. Disposable Isotonic gel EL507 electrodes were used to record EDA. Two 45cm BioNomadix Electro Leads were attached to electrodes measuring corrugator activity. Three 45cm BioNomadix Electro Leads were attached to electrodes measuring zygomatic activity and a ground electrode. An EDA Electrode Lead set was attached to electrodes measuring EDA. A BioNomadix Pulse Transducer and attached electrode was used to measure PPG. A BioNomadix Respiration Xdcr belt was used to measure respiration. All electrodes and leads used were manufactured by Biopac Systems Inc.

Materials

Digital film soundtracks were purchased through iTunes or copied from CDs by two research assistants. With the target moods in mind, one-minute excerpts were selected from each track for subsequent studies, including piloting and mood induction. For a full list of musical tracks and timecodes for excerpts used refer to Table 3. A 40-minute track was created for each musical excerpt chosen for the mood induction phase. During mood induction 40-minute stimulus tracks were presented through SONY MDR-XD100 stereo headphones. The first 20 minutes consisted of one of the 18 one-minute excerpts on a loop with one-second fades between each segment. The remaining 20 minutes contained a neutral, non-musical sound (ball drop) which occurred once a minute. The signal served as a cue for participants to fill out the subjective rating forms after the induction phase.

Procedure

Pilot Phase Eighteen excerpts from instrumental soundtracks were chosen using data that were collected in a two-stage pilot study. During the first stage, eight undergraduate students (aged 19-24) listened to 76 musical excerpts (refer to Table 3 for a full listing), with a break period of two minutes between excerpts, and rated them on valence, arousal,

tiredness, anxiety and sadness as well as familiarity, liking and predominance of the intended mood in the overall excerpts. Familiarity ratings were measured in order to eliminate tracks that may be too recognizable. Liking ratings were meant to assess the relationship between preference and ratings of pleasantness, but were not used in subsequent analyses. Predominance was measured to ensure that the mood perceived in the music was consistent throughout the entire excerpt.

Due to a lack of sufficient excerpts meeting the criteria for inclusion in phase I (see criteria below), a second stage of the pilot was created. In the second stage, 12 additional soundtrack excerpts (refer to Table 4 for full list) were presented using Qualtrics (a web based survey tool) in an online questionnaire mirroring the original pilot. The additional excerpts were selected based on recommendations made by colleagues and in online forums. In the second stage of the pilot, ratings were collected from 16 additional participants recruited online through social media.

Of all potential excerpts, 18 were chosen for the next phase of the study, with six tracks in each of the following mood conditions: sadness, fatigue and anxiety. Excerpts were chosen based on their discrete mood ratings (tiredness, sadness, and anxiety), given on a scale of 0-10 for each mood. The chosen excerpts were those with the highest mean rating in one of the mood categories, a significant main effect of the ANOVA comparing the mean ratings of the three different moods, and contrast values of less than .1. First, excerpts that had a significant main effect of mood rating were selected. Next, each qualifying excerpt was allocated into a condition based on which mean rating (tiredness, sadness, or anxiety) was the highest for that track. Lastly, tracks were required to have a contrast significance value of .1 or lower in order to be used in the next phase. If more than six excerpts qualified for a mood

condition, the six tracks with the highest mean ratings for the corresponding mood were retained.

Mood Induction Phase Refer to Figure 2 for a graphic representation of the experimental procedure. All participants were tested in a small quiet room located in the Ryerson Stress and Wellbeing Institute. Participants were randomly assigned to one of the 18 excerpts within one of the three mood induction conditions. First, participants were asked to complete a consent form. Then, participants were given the FSS, DAS22, Epworth Sleepiness Scale, the AIMS and a demographic information questionnaire. Once participants completed all forms they were notified that in the next phase of the study two belts were to be placed around their waist, three electrodes on their hand, and five electrodes on their face and neck.

While participants were standing, the wireless RSP receiver belt was placed loosely around the waist so that it would not constrict breathing. Next, the respiration belt was fastened around the waist while participants were instructed to take a deep breath in. Participants were then asked if the belt was tight enough so that they felt a bit of pressure while breathing in and out, but not so much that it made them uncomfortable, if they answered “no”, the belt was adjusted further, if they answered “yes” the next step of the procedure was performed.

Participants were asked to sit down and stretch out their non-dominant hand facing up. Handedness of participants was recorded in the experiment logbook. The palm of the hand and middle finger as well as the cheek, neck, and eyebrow area were cleaned with NuPrep skin prep gel. To ensure that skin was completely dry before attaching electrodes, equipment calibration was performed next. An AqKnowledge template was opened and,

following AqcKnowledge instructions, leads were attached to all receivers but not connected to any electrodes to perform software calibration.

Two EL507 electrodes were attached adjacent to each other on the outer pad region of the palm of the hand. The wireless EDA/PPG receiver was attached around the wrist and leads were attached to the electrodes (red lead was always attached to the front electrode to maintain consistency). A pulse transducer was attached to the middle finger.

Electromyographic electrodes (EL504) were always placed on the left side of the face. One EL504 electrode was attached to the back of the neck to serve as ground, approximately half way between the ear and the collarbone. Two EL504 electrodes were attached to the cheek over the zygomaticus major muscle, one approximately three quarters of an inch from the corner of the mouth, and the other adjacent to it and approximately half an inch higher (on a diagonal line). Two EL504 electrodes were attached to the forehead over the corrugator supercilii muscle, one above the inner corner of the eyebrow and the other adjacent to it. A wireless EMG receiver was attached to the upper arm. A 2-lead bundle was connected to the EMGa port and attached to corrugator electrodes. A 3-lead bundle was connected to the EMGb port and attached to the zygomatic electrodes as well as the lead (red leads were always attached to the medial electrode to maintain consistency, black lead was attached to ground electrode). Once participants were attached to all equipment a check was performed to ensure that all channels were functioning.

Next, participants were instructed to take a seat in a small quiet testing booth. A headphone set was placed over their ears and a sound level check was performed to adjust the volume of the music to a comfortable level. A sample of a few seconds of music was played and participants were asked if they would like the volume louder or quieter until a

comfortable volume was reached. Participants were then given a booklet containing all VAS mood scales and an additional page at the end of the booklet which asked for ratings of familiarity, predominance, and liking. Prior to induction participants filled out only the first page of the booklet, to serve as baseline for reported subjective mood. Once finished the first page of the booklet participants' physiological response was recorded at rest for 30 seconds, after which they heard a 40-minute stimulus track containing 20 minutes of music (induction phase) followed by 20 minutes of sound cues (postinduction phase).

During the 20-minute induction phase participants were instructed to only listen to the music while physiological response was being recorded. Once the music was finished, the postinduction phase began. During this phase participants filled out one page of the booklet each time they heard a non-musical audio cue. Participants were asked to flip each page over once finished and not to look back at their previous answers.

Once the testing stimulus played through, participants were escorted out of the testing booth, debriefed and disconnected from all equipment.

Data Analysis

Physiological Data Analysis. All data recorded with AcqKnowledge were visually inspected and issues with specific channels and participants were noted. Issues included missing or dropped signal as a result of electrode dislocation due to movement or sweating, partial trials as a result of uncompleted experiments, and excessive noise.

Once data were exported into Excel, all cases noted in visual inspection were examined and removed. Eight participants were removed entirely due to incomplete trials or equipment error. Five participants had unusable heart rate data. Five participants had unusable respiration data. Seven participants had unusable corrugator data. Eight participants

had unusable SCL data. Ten participants had unusable zygomaticus data. Z scores could not be computed for participant 63, corrugator; participant 67, corrugator, and participant 72, respiration. These channels were not used in the analysis.

Movement artifacts pose a problem in the analysis of physiological data. EMG recording may register movement as activity, generally resulting in a signal much stronger in magnitude than that of normal zygomaticus or corrugator activation. Movement can also cause noise in heart rate and respiration signals making meaningful quantification of rate difficult. Although this type of artifact does not generally pose an issue for SCL, excessive movement can result in displacement of electrodes, which can cause inaccurate estimates of SCL. The values observed when movement artifacts are present generally fall outside of the normal range. As a result, values over three standard deviations away from the mean were considered to be noise caused by movement artifacts, and were removed from the dataset.

Missing baseline values were left blank, and those channels were not used in the final analysis because scores could not be baseline corrected. When a single missing value was found in the data, the average value of the adjacent cells was computed, when two or more missing values were found, cells were left blank.

Electromyographic data were analyzed as raw values (Lang, Greenwald, Bradley & Hamm, 1993; Witvliet & Lang, 2007). Heart rate, SCL and respiration rate were normalized within participants and analyzed as z scores as in Sandstrom and Russo (2010).

Subjective Data Analysis Scores for the DASS21, FSS, AIMS, and ESS were computed using scoring keys (see appendix B) and entered into SPSS. VAS scores for valence and arousal were entered in SPSS as a score between 1-9. VAS scores for ratings of

tiredness, sadness, and anxiety, indicated on a 100 mm line, were measured using a ruler and a score between 1-100 was entered into SPSS.

Twenty participants did not complete the AIMS scale because it was introduced later in the study. One participant provided incomplete results for the DASS21 resulting in a missing depression sub-score. One participant did not complete the ESS. Fifteen participants did not complete the FSS. Seven participants did not provide liking and familiarity ratings. Twenty-eight participants either did not provide any predominance ratings or provided verbal responses. Two participants did not provide age and gender information. Five participants did not complete subjective evaluations following induction, 2 participants provided incomplete sadness, tiredness, and anxiety scores. One participant did not provide valence and arousal scores. Refer to Table 6 for summary of descriptive statistics. One-way ANOVAs were carried out to assess for between group differences. There were no significant differences on the DASSD, DASS A, DASS S, ESS, FSS or AIMS scales between mood conditions.

Results

AIMS

Scores on the AIMS scale were normally distributed, $D(107) = 0.07$, $p = .182$, with a mean of 112 and a standard deviation of 23.9, matching previously published results on the AIMS scale which reported a mean of 113 and a standard deviation of 23.8 (Sandstrom and Russo, 2011).

Subjective Data

First, subjective data were analyzed to establish the change in mood from baseline to immediately post induction using mixed-design repeated measures ANOVAs with time (baseline, immediately post induction) as a within-subjects independent variable, and

condition (sadness, fatigue, anxiety) as a between-subjects independent variable. Separate analyses were carried out with each of the subjective ratings as a dependent variable (valence, arousal, tiredness, sadness and anxiety). Summary of mean valence and arousal ratings and corresponding standard deviations can be seen in Table 7. Refer to Figure 3 for the location of excerpts in the study within the valence/arousal space based on participant ratings after induction.

For valence ratings, there was a significant main effect of time, $F(1, 76) = 42.5$, $p < .001$. partial $\eta^2 = .36$. The effect of condition and the time × condition interaction were not significant. Figure 4 depicts changes in valence ratings between baseline and ratings given immediately after induction in all conditions. As predicted, participants in all conditions reported a substantial decrease in valence after induction, however, the decrease was not significantly different between groups.

For arousal ratings, there was a significant main effect of time, $F(1, 74) = 15.5$, $p < .001$ partial $\eta^2 = .17$. The effect of condition and the time × condition interaction were not significant. As can be seen in Figure 5, arousal decreased in all conditions after induction.

There was a significant main effect of time for tiredness ratings, $F(1, 78) = 19.9$, $p < .001$ partial $\eta^2 = .203$. The effect of condition and the time × condition interaction were not significant. As can be seen in Figure 6, there was an increase in tiredness ratings after induction in all conditions.

No significant main effects or interactions were found for sadness ratings. There was no significant main effect for anxiety ratings, however the time × condition interaction was significant, $F(2, 78) = 3.17$, $p = .048$, partial $\eta^2 = .075$. Refer to Figure 7 for a depiction of anxiety rating changes between pre and post induction.

To assess the stability of affective change, indicating an induction of mood rather than a brief emotional response, ratings over the course of the postinduction period were evaluated next. ANOVAs were carried out once again for each subjective measure, with time as a within-subjects independent variable (21 levels: baseline and 20 minutes following induction, with one rating per minute, per subjective measure), and condition as a between subject independent variable.

For valence ratings, there was a significant main effect of time, $F(5.831, 329.699) = 9.6$, $p < .001$, partial $\eta^2 = .13$. Mauchly's test indicated that the assumption of sphericity had been violated for this measure for the main effect of time, $\chi^2(209) = 1123$, $p < .001$. Degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .258$). There was no significant effect of condition or condition × time interaction. As can be seen in Figure 8, ratings of valence remained stable after the initial postinduction reduction. The significant main effect of time was further explored using t-tests comparing (1) ratings at baseline with ratings post induction, and (2) ratings immediately post induction with ratings at the end of the post-induction phase. There was a significant reduction in valence in all conditions between pre and post induction, $t(78)=6.47$, $p<.001$, and no significant change in ratings between the beginning and end of the post induction phase, supporting the stability of mood change.

For arousal ratings, the main effect of time was significant, $F(4.467, 250.141) = 3.3$, $p = .008$, partial $\eta^2 = .056$. Mauchly's test indicated that the assumption of sphericity had been violated for this measure for the main effect of time, $\chi^2(209) = 1157$, $p < .001$. Degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .223$). The time × condition interaction failed to remain significant following a Greenhouse-Geisser

correction, and the main effect of condition was not significant. As depicted in Figure 9, it is evident by visual inspection that arousal ratings began to converge after the first five minutes of the postinduction period. The significant main effect of time was further explored using t-tests comparing (1) ratings at baseline with ratings post induction, and (2) ratings immediately following induction with ratings at the end of the postinduction phase. There was a significant reduction in arousal irrespective of condition between pre and post induction, $t(76) = 4.0$, $p < .001$, and no significant difference in ratings between the beginning and end of the post induction phase, once again supporting the stability of mood change.

For tiredness ratings, the main effect of time was significant, $F(5.529, 370.465) = 3.4$, $p = .004$, partial $\eta^2 = .048$. Mauchly's test indicated that the assumption of sphericity had been violated for the main effect of time for this measure, $\chi^2(209) = 1180$, $p < .001$. Degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .276$). As shown in Figure 10, all conditions exhibited an increase in tiredness ratings post induction, and this effect was strongest in the fatigue condition. All conditions also displayed a return to baseline following the induction. The significant main effect of time was further explored using t-tests comparing (1) ratings at baseline with ratings post induction, and (2) ratings immediately post induction with ratings at the end of the post-induction phase. An increase in tiredness ratings between pre and post induction was significant across groups $t(80) = 4.49$, $p < .001$. The decrease in tiredness ratings during the post induction phase was also significant, $t(73) = 2.12$, $p = .005$, indicating that the increase in tiredness caused by mood induction was not stable.

Sadness ratings did not yield any significant effects. For anxiety ratings, there were no significant main effects for either time or condition, however, there was a significant condition × time interaction, $F(6.095, 204.168) = 2.18$, $p = .045$, partial $\eta^2 = .061$. Mauchly's test indicated that the assumption of sphericity had been violated for the main effect of time for this measure, $\chi^2(209) = 1733$, $p < .001$. Degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .152$). Levene's test indicated that the assumption of homogeneity of error variances had been violated in two of the 21 time points (T15, $F(2, 67) = 4.41$, $p = .016$, and T16, $F(2, 67) = 3.87$, $p = .026$). As shown in Figure 11, anxiety ratings showed little fluctuation in the sadness and anxiety conditions, but revealed a different pattern in the fatigue condition. Based on visual inspection of the ratings, after an initial decrease in anxiety, ratings began to climb following the first five minutes of the postinduction period and returned to baseline by the end of the 20 minutes.

In summary, all conditions exhibited significant reductions in valence and arousal during the mood induction phase, which remained stable during the postinduction phase. All conditions exhibited a significant increase in tiredness ratings immediately after the induction phase, however, ratings returned to baseline by the end of the postinduction phase. No significant main effects were evident for ratings of sadness and anxiety. Finally, anxiety ratings yielded a significant condition × time interaction in both the induction and postinduction phases, with the fatigue condition exhibiting a reduction in ratings immediately after the induction phase, which returned to baseline ratings by the end of the postinduction phase.

Strongest changes in target mood caused by induction (by individual track)

To establish the best excerpt for each induction, target mood ratings within each induction condition were analyzed using paired sample t-tests. Refer to figures 12-14 for a depiction of target mood changes between pre and post induction in each track by condition.

Of the six excerpts in the sadness condition, four caused an increase in sadness ratings. Paired-sample t-test were conducted for each excerpt. No significant changes in sadness ratings from baseline were found for any of the excerpts. However, because samples for each excerpt were small, effect sizes were calculated to further investigate the magnitude of change. For the two excerpts with the greatest difference between pre and post induction, effect sizes were $r = .69$ (excerpt S39), and $r = .7$ (excerpt S40). I expect that with larger sample size, these tracks are likely to reach significance. Power analysis using G*Power revealed that significant results can be expected with a sample size of 30 participants per excerpt condition for S39 and S40 with power set at .95. The other two excerpts that caused an increase in sadness ratings had effect sizes of $r = .38$ (S13), and $r = .42$ (S25). The two excerpts that caused a decrease in sadness ratings, counter to prediction, had effect sizes of $r = .55$ (S16), and $r = .11$ (S52). Refer to Figure 12 for a graphic representation of change in sadness ratings by excerpt in the sadness condition.

Tiredness ratings in all six excerpts in the fatigue condition increased from baseline, but none of the findings reached significance. Effect sizes were calculated for each excerpt. Strongest effect sized were $r = .61$ (S37), $r = .59$ (S19), and $r = .55$ (S11). Power analysis using G*Power revealed that with power set at .95, sample sizes of 37, 40 and 45 participants would be required for excerpts S11, S19, and S37, respectively. Other excerpts achieved

effect sizes of r = .44 (S35), r = .42 (S25), and r = .20 (SP7). Refer to Figure 13 for a graphic representation of change in tiredness ratings by excerpt in the fatigue condition.

Of the six tracks in the Anxiety condition, two caused increases in anxiety ratings. Excerpt S34 caused a significant increase in anxiety ratings $t(5)$ = -.3.099, p = .027, r = .81. Although S47 failed to reach significance (p = .055), the calculated effect size was high, r = .87. Power analysis using G*Power indicated that a sample size of 20 participants was required for excerpt S47 with power set at .95. For excerpts which caused mood changes in the opposite than predicted direction effect sizes were r = .067 (S27), r = .027 (S44), r = .74 (S23), and r = .4 (S50). Refer to Figure 14 for visualization of anxiety rating changes by excerpt between pre and post induction in the anxiety condition.

In summary, two to three tracks in each mood condition yielded changes in mood ratings in the predicted direction during the induction phase, with substantial effect sizes. Only one excerpt, in the anxiety condition, reached significance. Graphing the best two tracks in each condition within Russell's circumplex based on valence and arousal ratings for each track (see Figure 15), it can be seen that in contrast with Figure 4, the tracks occupy more distinct clusters within the model.

Physiological data

Figures 16-19 show how physiological response unfolded over the course of the 20-minute induction period, with each time point representing the signal averaged in 20-second segments. Mixed design repeated measures ANOVAs were carried out for each of the five physiological measures, with time as a within-subjects independent variable (60 levels, with each level representing an average over 20 seconds) and condition (fatigue, sadness, anxiety)

as a between-subjects variable. There were no significant main effects or interactions for heart rate, corrugator or zygomatic activity.

For respiration rate, the ANOVA indicated a significant main effect of time, $F(19.875, 1252.135) = 2.36$, $p = .001$, partial $\eta^2 = .036$. Mauchly's test indicated that the assumption of sphericity had been violated for this measure, $\chi^2(1769) = 2863$, $p < .001$ for the main effect of time. Degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .337$). As can be seen in Figure 17, respiration rate increased in all conditions during induction.

For SCL, the ANOVA revealed a significant main effect of time, $F(10.396, 675.752) = 2.47$, $p = .006$, partial $\eta^2 = .037$ and a significant time × condition interaction, $F(20.792, 675.752) = 1.83$, $p = .014$, partial $\eta^2 = .053$. Mauchly's test indicated that the assumption of sphericity had been violated for the main effect of time for this measure, $\chi^2(1769) = 4281$, $p < .001$. Degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .176$). As can be seen in Figure 18, during the first 10 minutes of mood induction, SCL showed a decrease in the fatigue and anxiety conditions while the sadness condition showed an increase. Based on visual inspection, it is evident that trends converged between the 10th and 15th minutes, followed by a gradual drop for the fatigue condition only.

To further explore the physiological responses, data were reanalyzed using a moving time window of four minutes. Repeated measures ANOVAs were conducted on each time window with one of the five physiological features as a dependent variable, with condition (sadness, fatigue, anxiety) as a between-subjects independent variable, and time (12 timespoints, each 20 seconds in duration) as a within-subjects independent variable. Table 8 shows the summary of results for between-subjects effects of condition for all ANOVAs. As

may be seen in the table, the biggest group differences were observed between three and 10 minutes.

SCL was the only measure that yielded significant results for main effect of condition, these differences were evident in the 3-6 minute window, $F(2, 67) = 3.44$, $p = .038$, partial $\eta^2 = .093$; 5-8 minute window, $F(2, 67) = 4.28$, $p = .018$, partial $\eta^2 = .113$; and 7-10 minute window, $F(2, 67) = 4.13$, $p = .020$, partial $\eta^2 = .11$ As hypothesized, the fatigue condition showed the lowest levels of skin conductance, however, the anxiety condition failed to exhibit levels higher than the sadness condition. Based on Field's (2009) recommendation, since group sizes were unequal, Hochberg's GT2 post hoc tests were used to examined between group effects. Hochberg's post hoc tests revealed that the fatigue condition had significantly lower SCL than the sadness condition in the 3-6 minute window, $p = .032$, 5-8 minute window, $p = .014$, and 7-10 minute window, $p = .017$. In all time windows, the sadness condition did not differ significantly from the anxiety condition, and the anxiety condition did not differ significantly from either.

For SCL, Levene's test indicated that the assumption of homogeneity of variance had been violated for six of the 30 time points within the three to 10 minute range, a summary of significant Levene's test results can be seen in Table 12. Transformations were attempted to correct for the violation of the assumption of homogeneity of variance (Square and Cube transformations, others could not be computed because of negative values in the dataset), but homogeneity of variance could not be achieved.

As a result, each time point was evaluated using the Kruskall-Wallis. Due to the resulting number of comparisons, group differences did not reach significance in any of the time points after a Bonferroni correction was applied (.05/23 = .002).

Gastwirth, Gel and Miao (2009) argue that Levene's test tends to be inflated when group sizes are unequal, as in the present sample. It is also worth noting that the assumption was met for most levels of the time variable. A continuation of this study following the thesis will attempt to correct for this violation by equating group sizes.

Main effects of time and time × condition interactions were found to be significant in several 4-minute time windows for measures of SCL, heart rate and respiration rate. For SCL, the ANOVA revealed a significant main effect of time in the 1-4 minute window, $F(5.12, 342.92) = 2.82$, $p = .045$, partial $\eta^2 = .033$, as well as a significant time × condition interaction, $F(1.57, 342.92) = 2.07$, $p = .025$, partial $\eta^2 = .058$. Mauchly's test indicated that the assumption of sphericity had been violated for SCL in this time window, $\chi^2(65) = 251$, $p < .001$ for the main effect of time. Degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .465$). As shown in Figure 18, the fatigue and anxiety conditions exhibited a decrease in SCL, while the sadness condition showed an increase within this time window.

For SCL, the ANOVA revealed a significant main effect of time in the 7-10 minute window, $F(12.06, 403.94) = 2.18$, $p = .044$, partial $\eta^2 = .031$. Mauchly's test indicated that the assumption of sphericity had been violated for skin conductance in this time window minute time window, $\chi^2(65) = 213.89$, $p < .001$ for the main effect of time. Degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .548$). As can be seen in Figure 18, the fatigue condition exhibited a decrease while the anxiety and sadness conditions showed an increase within this time window.

For SCL, the ANOVA revealed a significant condition × time interaction effect in the 9-12 minute window, $F(12.99, 435.146) = 2.75$, $p = .001$, partial $\eta^2 = .076$. Mauchly's test

indicated that the assumption of sphericity had been violated for SCL in this time window, $\chi^2(65) = 192$, $p < .001$ for the main effect of time. Degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .590$). As shown in figure 18, in this time window the fatigue condition showed an increase, the sadness condition showed a decrease, and the anxiety condition exhibited fluctuations.

For SCL, the ANOVA revealed a significant condition × time interaction effect in the 11-14 minute window, $F(13.68, 451.48) = 2.48$, $p = .002$, partial $\eta^2 = .07$. Mauchly's test indicated that the assumption of sphericity had been violated for SCL in this time window, $\chi^2(65) = 206$, $p < .001$ for the main effect of time. Degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .622$). As shown in Figure 18, the fatigue condition showed and increase, while the sadness and anxiety condition showed a decrease.

For SCL, the ANOVA revealed a significant condition × time interaction effect in the 15-18 minute time window, $F(13.73, 453.15) = 2.46$, $p = .002$, partial $\eta^2 = .07$. Mauchly's test indicated that the assumption of sphericity had been violated for SCL in the 15-18 minute window, $\chi^2(65) = 213$, $p < .001$ for the main effect of time. Degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .624$). As can be seen in Figure 13, the sadness condition stayed relatively stable, while the anxiety condition showed an increase and the sadness condition showed a decrease.

For heart rate, the ANOVA revealed a significant main effect of time in the 1-4 minute window, $F(7.09, 460.97) = 2.22$, $p = .03$, partial $\eta^2 = .033$. Mauchly's test indicated that the assumption of sphericity had been violated for heart rate in this time window, $\chi^2(65) = 177$, $p < .001$ for the main effect of time. Degrees of freedom were corrected using

Greenhouse-Geisser estimates of sphericity (ε = .645). As can be see in figure 16, all groups showed an initial decrease in heart rate, followed by an increase.

For respiration rate, the ANOVA revealed a significant time × condition interaction in the 11-14 minute window, $F(17.65, 573.64) = 2.00$ $p = .09$, partial $\eta^2 = .028$. Mauchly's test indicated that the assumption of sphericity had been violated for respiration in this time window, $\chi^2(65) = 87.47$, $p<.001$ for the main effect of time. Degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (ε = .802). Within this timeframe the sadness condition remained relatively stable while the fatigue and sadness condition showed fluctuated.

For respiration rate, the ANOVA revealed a significant time × condition interaction in the 13-16 minute window, $F(1.26, 562.234) = 1.84$, $p = .011$, partial $\eta^2 = .053$. Mauchly's test indicated that the assumption of sphericity had been violated for respiration rate in this time window, $\chi^2(65) = 101$, $p < .01$ for the main effect of time. Degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (ε = .774). As can be seen in Figure 17, respiration fluctuated in all conditions within this timeframe.

For respiration rate, the ANOVA revealed a significant time × condition interaction in the 15-18 minute window, $F(17.17, 557.913) = 2.45$, $p = .001$, partial $\eta^2 = .07$. Mauchly's test indicated that the assumption of sphericity had been violated for respiration rate in this time window, $\chi^2(65) = 100$, $p < .01$ for the main effect of time. Degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (ε = .780). As can be seen in Figure 17, respiration fluctuated in all conditions within this timeframe.

Finally, mixed design repeated measures ANOVAs were used to assess physiological changes post induction. No significant main effects or interactions were found.

In summary, respiration showed a significant main effect of time during the mood induction period, with an increase across all conditions. SCL also yielded a significant main effect of time in the mood induction phase. From visual inspection it can be speculated that this was driven by the reduction in SCL in the fatigue and anxiety conditions. Further analyses revealed that the fatigue condition showed significantly lower levels of SCL than the sad, but the not anxious condition in the 3-6 minute window, 5-8 minute window, and 7-10 minute window of the mood induction phase.

Questionnaires and Subjective and Physiological Measures

Kolmogorov-Smirnov tests revealed that the assumption of normality was violated for several questionnaire measures. Field (2009) states that the assumption of normality can be easily violated when sample size is large and that, when comparing groups, the distribution within groups is more meaningful that the overall distribution. Consequently, normality was evaluated within each condition. Epworth scores were significantly non-normal in the fatigue condition, $D(33) = 0.20$, $p = .002$. FSS scores were significantly non-normal in the sadness condition, $D(39) = 0.16$, $p = .015$. DASS depression score were significantly non-normal in the sadness, $D(39) = 0.17$, $p = ,006$, fatigue, $D(33) = 0.21$, $p = .001$, and anxiety, $D(37) = 0.15$, $p = .034$, conditions. DASS anxiety scores were significantly non-normal in the sadness, $D(39) = 0.17$, $p = .006$, fatigue, $D(33) = 0.24$, $p < .001$, and anxiety, $D(37) = 0.18$, $p = .005$, conditions. DASS stress scores were significantly non-normal in the sadness condition, $D(39)= 0.22$, $p < .001$. Transformation of the data was attempted to correct for these violations of normality but normality could not be achieved.

Multiple regressions were used examine the effects of depression, anxiety, stress, sleepiness and fatigue scores (using the DASS21, FSS, and ESS) on changes from baseline to

postinduction mood ratings across participants in all conditions. For subjective ratings, change scores were computed for each mood rating, and used as dependent variables, questionnaire scores were entered using the backward method as independent variables. None of the questionnaire scores contributed significantly to mood change.

The pattern of physiological changes caused by mood induction indicated that participants showed strongest group difference between 3 and 10 minutes following which group effects diminished, likely indicating habituation to the procedure. Based on these findings, physiological changes within the four-minute time window that indicated the strongest group differences between three and ten minutes (Table 8) were used to assess the effects of depression, anxiety, stress, sleepiness, and fatigue. There was a significant contribution of FSS scores for heart rate in the five and eight minute window (refer to Table 9 for complete regression summary), and a significant effect of depression scores on the DASS21 for SCL in the five and eight minute window (refer to Table 10 for complete regression summary). For zygomatic activity in the seven and 10 minute window, depression, anxiety and stress score were retained in the final model, however only the contribution of anxiety scores was significant (refer to table 11 for complete regression summary).

Discussion

As predicted, participants in all conditions reported a decrease in valence ratings, although, contradictory to the hypothesis, the change was similar across groups. This reduction was sustained through the entire postinduction phase indicating that the change in the valence component of the mood state was stable.

All groups showed a decrease in arousal ratings, and paralleling this finding, an increase in tiredness ratings. Although I hypothesized that the anxiety condition would

experience an increase in arousal while the sadness and fatigue conditions experience decreases, a decrease in all conditions is not surprising. This observed drop in arousal is likely the result of students sitting still without having to engage in any tasks for the duration of induction. Although the time × condition interaction for arousal data were not significant, the trends note by visual inspection suggest that, as hypothesized, the anxiety condition reported the smallest reduction in arousal, followed by the sadness condition, and the fatigue condition reported the greatest reduction. As predicted, the fatigue condition showed the highest tiredness and lowest arousal ratings post induction. Although arousal ratings began to increase during the postinduction phase, the change in arousal from the beginning to the end of the postinduction phase was not significant, indicating a degree of stability in arousal change caused by the mood induction. Tiredness data showed a trend echoing the arousal ratings, with initial between-group differences disappearing after roughly five minutes. These trends support the notion that the reduction in arousal was significantly affected by the lack of activity during the induction phase as arousal ratings increased once participants were required to engage in a task.

Contrary to my hypothesis, no significant increase in anxiety ratings was found in the anxiety condition. However, the significant time × condition interaction revealed that the fatigue condition exhibited a decrease in anxiety ratings, likely indicating a calming effect of music in the fatigue condition. It is worth noting that this reduction in anxiety ratings was not stable after the end of the induction period, as anxiety ratings returned to baseline before the end of the postinduction phase.

The fatigue condition exhibited the lowest heart rate, followed by the sadness condition, and the anxiety condition experienced the highest heart rate; however the main

effect of condition failed to reach significance. As predicted, the fatigue condition exhibited the lowest SCL, however, contrary to prediction, SCL was higher in the sadness than the anxiety condition. None of the hypotheses regarding respiration or electromyographic activity were supported.

The complex physiological findings in this study call for further research. Firstly, the present study would benefit from equal group sizes across conditions. This could help resolve issues of normality in the data and increase power. A lack of power and unequal group sizes could have diminished some of the effects, such as the group differences in heart rate and arousal ratings.

A control condition that does not experience any kind of mood induction is advisable in future. Because the reduction in arousal and increase in tiredness ratings in the study may be attributed to this lack of activity, the inclusion of a control group that experiences the same procedure without listening to music may help interpret findings. A control group could provide a more meaningful reference point for establishing whether the contribution of music was substantial. The inclusion of a control group may also help elucidate the mixed physiological findings by showing what physiological activity under similar conditions when no stimulus is present is like.

In addition, more sophisticated means of cleaning up noise in the data are required. Because it is impossible to avoid movement artifacts, especially over a period of 20 minutes or longer, a procedure which includes removing noisy data and interpolating the gaps is suggested for future analyses to obtain a more accurate account of physiological response. Lastly, future studies should include behavioural tasks in addition to subjective and physiological measures to evaluate the validity of the method.

The mixed findings in the sadness induction literature (Kreibig, 2010) emphasize the importance of differentiating between activating and deactivating sadness in future studies. Since no attempt to differentiate these two types of sadness was made in this study, the sadness condition could have potentially resulted in a combination of both. This could be problematic because activating sadness can cause an increase in physiological arousal, while deactivating sadness can result in the opposite. A mix of both types of sadness in the condition can increase the response variability within the group, making meaningful interpretation of results more challenging. Moreover, during the pilot phase tracks were chosen in part based on their distinctiveness from the other two mood categories, and differentiation between the sadness and fatigue conditions was particularly difficult. As a result, the tracks chosen for the sadness condition were likely those which were more activating in nature, since these were less similar to the tired excerpts. This could potentially explain high SCL in the sadness condition, which were more similar to the anxiety condition than the fatigue condition response. In order to achieve better specificity in mood induction, the activating and deactivating subtypes of sadness need to be differentiated in future studies.

Another possible explanation is that the tracks used in the study were not effective for the procedure. This is particularly problematic in light of findings that suggest that perceived emotions may be different from felt emotions elicited by the same music. More extensive validation methods prior to induction may be required to choose suitable tracks. This could be done by increasing the sample of raters in the pilot phase of future validation studies, or consulting music experts, such as musicologists or composers, as has been done in some previous studies. Ideally, a combination of these methods would be performed.

Finally, it is important to consider that response to music is highly variable among individuals and largely affected by personal characteristics such as preference, associations, and memories (Juslin & Västfjäll, 2008; Eerola and Vuoskoski, 2013). Self-selected music has only been used in a handful of studies because it limits experimenter control. However, findings illustrate that self-selected music may be more appropriate for mood mediation. Participants show highest enjoyment, reduction in tension-anxiety and improved driving performance when listening to self-selected music, whereas this pattern was reversed for high arousal experimenter selected music (Cassidy & MacDonald, 2009). Self-selected music is also often used in therapeutic context to alleviate distress and physical symptoms (Clark, Isaacks-Downton, Wells, Redlin-Frazier, Eck, Hepworth, & Chakravarthy, 2006) and has been found to be more effective at relieving pain (Mitchell & MacDonlad, 2006). The availability of tools such as the MIR toolbox, which allow in-depth analysis of audio features, can help increase experimenter control through the utilization of audio-feature based matching between participants (Vuoskoski & Eerola, 2013. Future studies can expand upon musical mood induction work by matching tracks on audio features while allowing participants to choose their own music. Such work could potentially provide more effective and valid means of inducing mood. Furthermore, as a result of the augmented effects of self-selected music on mood changes, this method may be a stronger candidate for inducing highly specific moods. This is a promising avenue in light of the present study's findings, which illustrate the challenges of inducing specific moods through music.

Appendix A

The role of audio features in mood induction

Several musical features have been investigated in the context of emotional perception and experience in music. Mode (Hevner, 1937; Husain, Thompson & Schellenberg, 2002) and tempo (Hevner, 1937; Husain, Thompson & Schellenberg, 2002; Bernardi, Porta & Sleight, 2006) have been linked to both perceptual and physiological experiences of emotions. The role of pitch in emotional perception has been investigated as well, though attained less concrete findings (Hevner, 1937; Ilie & Thompson, 2006; Schellenberg, Krysciak & Campbell, 2000). In comparison, features pertaining to timbre and dynamics have received much less attention in the literature.

The earliest series of studies on the effects of musical features on emotion were conducted by Kate Hevner in the 1930s. By manipulating certain aspects of music compositions, such as mode and tempo, while holding all other variables constant, she was able to isolate the effects of these variables on emotion perception. Emotions were assessed with adjective checklists which participants filled out while listening to music. Findings revealed that minor mode was most strongly associated with sad music, and major mode with graceful and happy music. Slow tempo resulted in participants selecting more dreamy and serene adjectives, and fast tempo in happy and exciting ones. Further investigation of mode and tempo by Husain, Thompson and Schellenberg (2002) revealed that major mode yielded significantly higher mood ratings than the minor mode, and that fast tempi resulted in significantly higher arousal ratings than slower tempi. Physiologically, heart rate has been reported to be higher when listening to major than minor mode (Gomez & Danuser, 2007). Faster tempi have been shown to significantly increase ventilation, breathing rate, systolic,

and diastolic blood pressure, and mid-cerebral artery blood flow (Bernardi, Porta & Sleight, 2006) as well as SCLs and heart rate (Gomez & Danuser, 2007).

Pitch has been investigated in relation to expression of emotion in music as well, however findings on this feature are less clear. In Hevner's work, low pitch was associated with sad music, and high pitch with graceful music, though the latter relationship was less strong (Hevner, 1937). In contrast, Ilie and Thompson found that low-pitched music was rated as more pleasing than high-pitched music (2006). Schellenberg, Krysciak and Campbell (2000) manipulated rhythm and pitch in happy, sad and scary melodies, and asked participants to rate how accurately each version conveyed the intended content. Participants' ratings were affected more by pitch than rhythm variations. However, the effects of pitch and rhythm were different across the three mood conditions, indicating that the relationship is not straightforward.

Timbre is the psychoacoustic quality of sound that distinguishes different musical instruments and voices even when they are producing the same note at the same loudness. Timbre is spread throughout the entire signal, making it possible to identify even when only a brief segment of the sound is present (Moore, 1995). Musical excerpts as short as 250 ms can be accurately judged on emotional content (Peretz, Gagnon & Bouchard, 1998). Since these excerpts are not long enough to contain tempo or mode information, timbre must play a key role in the expression of emotions in music (Eerola, Alluri & Ferrer, 2008). This finding extends to judgments of emotions in isolated musical sounds. A high level of agreement is evident in the ratings of participants when asked to discern emotional content of brief instrument sound samples. Moreover, timbral characteristics are found to be predictive of emotional judgments. (Eerola, Alluri & Ferrer, 2008). Hailstone et al (2009) conducted an

experiment in which novel compositions written and rated as sad, happy, angry and fearful were played using different instruments in the same mode and tempo. A significant interaction between timbre and emotion judgment was identified, meaning that certain moods were judged more accurately when played in some timbres than others.

Dynamics, which refers to the intensity of the sound signal, has been linked to emotional expressiveness and perception. In a study by Kamenetsky, Hill, and Trehub, (1997) participants listened to 30-second music excerpts, some of which varied in dynamics and others did not. When asked to rate the excerpts on likeability and emotional expressiveness, participants provided higher ratings on both scales for the excerpts that varied in dynamics. In Ilie and Thompson's (2006) study, participants rated low-intensity music higher on valence, and high-intensity music higher on tension arousal and energy arousal.

These findings suggest that audio and music features play an important role in the perception of emotions in music. Moreover, studies show that audio features can be used to predict emotional ratings of valence and arousal (Coutinho & Cangelosi, 2011; Yang, Lin, Su & Chen, 2007; Schubert, 2004) as well as continuous ratings of some basic emotions (Eerola, Lartillot & Toiviainen, 2009). In order to examine this relationship more extensively, Study II analyzed audio and music features in conjunction with subjective response ratings obtained in Study I. This analysis of audio and music features as predictors of affective response was largely exploratory, however, I expected that a relationship between the extracted features and affective response would arise.

Method

All audio tracks were analyzed using MIRtoolbox and custom scripts in MATLAB (Lartillot, 2013; Lartillot, Eerola, Toiviainen & Fornari, 2008). Twelve low-level audio features were extracted from each track and represented as numerical values averaged across the entire duration of the excerpts. Table 13 provides a summary and description of the 12 audio features used in the analysis.

Results

One-Way ANOVAs were carried out to assess differences between audio features across the three conditions. 12 audio features served as dependent variables (root mean square (RMS), low energy, event density, tempo, pulse clarity, zero cross, centroid, spread, rolloff, brightness, irregularity, and mode).

Levene's test indicated that the assumption of homogeneity of variance was violated for tempo, pulse clarity, spread, rolloff and brightness. These measures were then analyzed using the Kruskall-Wallis, a non-parametric equivalent of the One-Way ANOVA test (Field, 2009).

The One-Way ANOVA revealed a significant effect of condition for RMS, $F(2, 15) = 17.4$, $p < .001$, $\omega^2 = .61$. Post hoc analysis using Tukey's HSD indicated that RMS was similar in the sadness and fatigue conditions, and significantly higher in the anxiety condition ($p < .001$).

The Kruskall-Wallis revealed a significant effect of condition for pulse clarity, $H(2) = 11.6$, $p = .003$. Mann-Whitney tests were used as follow up, with a Bonferroni adjustment for multiple comparisons (.05/3 = .017). Pulse clarity was significantly higher in the fatigue ($Mdn = .187$) than in the sadness condition ($Mdn = .063$), $U = 1.00$, $z = -2.72$, $p = .006$, $r =$ -

.78. Pulse clarity was also significantly higher in the anxiety (*Mdn* = .418) than in the sadness (*Mdn* = .063) condition, $U = 0.00$, $z = -2.89$, $p = .004$, $r = .83$. There was no significant difference between the fatigue and anxiety conditions.

The Kruskall-Wallis also revealed a significant effect of condition for spread, $H(2) = 8.56$, $p = .014$. Mann-Whitney tests were used as follow up, with a Bonferroni adjustment for multiple comparisons (.05/3 = .017). Spread was significantly higher in in the fatigue (*Mdn* = 1.92) than the anxiety (*Mdn* = .83) condition, $U = 0.00$, $z = -2.82$, $p = .004$, $r = .81$.

Based on the results of the ANOVA and the Kruskall-Wallis tests, which identified that RMS, pulse clarity, and spread were significantly different between the three mood conditions, multiple regressions were used to assess the contribution of these features to subjective measures of mood. Difference scores were calculated for the subjective ratings of participants between pre and post induction on all five subjective measures: valence, arousal, tiredness, sadness and anxiety. Multiple regressions were then carried out with mood change scores for each subjective evaluation as dependent variables, and RMS, pulse clarity and spread as predictor variables. Since this analysis was somewhat exploratory and there was no reason to believe that all three variables would significantly predict subjective ratings, a stepwise regression was used to remove predictors that did not make meaningful contributions to the model. Based on a recommendation by Field (2009) a 'backward' was chosen over the 'forward' method, which is more susceptible to suppressor effects.

There was a significant negative relationship between spectral spread and arousal and anxiety ratings. As spectral spread increased, arousal and anxiety ratings decreased. The relationship was stronger for anxiety than arousal ratings. Findings are summarized in Table 14 and 15.

Discussion

As predicted, the three mood conditions exhibited different audio feature profiles, more specifically, differences in RMS, pulse clarity and spectral spread. Spread contributed significantly to arousal and anxiety ratings. Regression models indicated that spectral spread (bandwidth of frequency spectrum) accounted for the largest proportion of variance in arousal and anxiety ratings. As spectral spread increased, arousal and anxiety ratings decreased. Higher spread values were associated with lower anxiety and arousal ratings, meaning that tracks that showed more variability in the frequency spectrum were less arousing, and less anxiety provoking. These findings need to be corroborated in future work through an identification of novel excerpts that vary in arousal and anxiousness based on their spectral spread characteristics.

Appendix B

DASS 21 NAME _____ DATE _____

Please read each statement and circle a number 0, 1, 2 or 3 which indicates how much the statement applied to you over the past week. There are no right or wrong answers. Do not spend too much time on any statement.

The rating scale is as follows:

0 Did not apply to me at all - NEVER

1 Applied to me to some degree, or some of the time - SOMETIMES

2 Applied to me to a considerable degree, or a good part of time - OFTEN

3 Applied to me very much, or most of the time - ALMOST ALWAYS

		N	S	O	AA	FOR OFFICE USE D	A	S
1	I found it hard to wind down	0	1	2	3			
2	I was aware of dryness of my mouth	0	1	2	3			
3	I couldn't seem to experience any positive feeling at all	0	1	2	3			
4	I experienced breathing difficulty (eg, excessively rapid breathing, breathlessness in the absence of physical exertion)	0	1	2	3			
5	I found it difficult to work up the initiative to do things	0	1	2	3			
6	I tended to over-react to situations	0	1	2	3			
7	I experienced trembling (eg, in the hands)	0	1	2	3			
8	I felt that I was using a lot of nervous energy	0	1	2	3			
9	I was worried about situations in which I might panic and make a fool of myself	0	1	2	3			
10	I felt that I had nothing to look forward to	0	1	2	3			
11	I found myself getting agitated	0	1	2	3			
12	I found it difficult to relax	0	1	2	3			
13	I felt down-hearted and blue	0	1	2	3			
14	I was intolerant of anything that kept me from getting on with what I was doing	0	1	2	3			
15	I felt I was close to panic	0	1	2	3			
16	I was unable to become enthusiastic about anything	0	1	2	3			
17	I felt I wasn't worth much as a person	0	1	2	3			
18	I felt that I was rather touchy	0	1	2	3			
19	I was aware of the action of my heart in the absence of physicalexertion (eg, sense of heart rate increase, heart missing a beat)	0	1	2	3			
20	I felt scared without any good reason	0	1	2	3			
21	I felt that life was meaningless	0	1	2	3			
					TOTALS			

References

Albersnagel, F. A. (1988). Velten and musical mood induction procedures: a comparison with accessibility of thought associations. *Behaviour research and therapy*, *26*(1), 79–96.

Andrews, A. J., Nespoli, G., and Russo, F. A. (2011). FeatureFinder (Version 2.2). Available online at: http://www.featurefinder.ca

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & Van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. *Psychological Bulletin*, *133*(1)*,* 1-24.

Bernardi, L., Porta, C., & Sleight, P. (2006). Cardiovascular, cerebrovascular, and respiratory changes induced by different types of music in musicians and non-musicians: the importance of silence. *Heart (British Cardiac Society)*, *92*(4), 445-52. doi:10.1136/hrt.2005.064600

Blechert, J., Lajtman, M., Michael, T., Margraf, J., &Wilhelm, F. H. (2006). Identifying anxiety states using broad sampling and advanced processing of peripheral physiological information. *Biomedical Sciences Instrumentation*, *42*, 136–141.

Bower, G. H., Forgas, J. P. (2001). Mood and Social Memory. In Forgas (Eds.). *Handbook of Affect and Social Cognition* (pp. 97-121). Taylor and Francis e-Library, 2009.

Bradley, M., & Lang, P. J. (1994). Measuring Emotion: The Self-Assessment Manikin and the Semantic Differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(1), 49-59.

Cacioppo, J. T., Tassinary, L. G., Berntson, G. (2007). Handbook of Psychophysiology. New York: Oxford University Press.

Carney, C. E., Moss, T. G., Lachowski, A. M., & Atwood, M. E. (2013). Understanding Mental and Physical Fatigue Complaints in Those With Depression and Insomnia. *Behavioral sleep medicine*, (October 2013), 37–41. doi:10.1080/15402002.2013.801345

Cassidy, G., & Macdonald, R. (2009). The effects of music choice on task performance: A study of the impact of self-selected and experimenter-selected music on driving game performance and experience. *Musicae Scientiae*, *13*(2), 357–386. doi:10.1177/102986490901300207

Chaudhuri, a, & Behan, P. O. (2000). Fatigue and basal ganglia. *Journal of the neurological sciences*, *179*(S 1-2), 34–42. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11054483

Chepenik, L. G., Cornew, L. a, & Farah, M. J. (2007). The influence of sad mood on cognition. *Emotion (Washington, D.C.)*, *7*(4), 802–11. doi:10.1037/1528-3542.7.4.802

Clark, D. M. (1983). On the induction of depressed mood in the laboratory: Evaluation and comparison of the velten and musical procedures. *Advances in Behaviour Research and Therapy*, *5*(1), 27-49. doi:10.1016/0146-6402(83)90014-0

Clark, M., Isaacks-Downton, G., Wells, N., Redlin-Frazier, S., Eck, C., Hepworth, J. T., & Chakravarthy, B. (2006). Use of preferred music to reduce emotional distress and symptom activity during radiation therapy. *Journal of music therapy*, *43*(3), 247–65. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17037953

Coutinho, E., & Cangelosi, A. (2011). Musical emotions: predicting second-by-second subjective feelings of emotion from low-level psychoacoustic features and physiological measurements. *Emotion*, *11*(4), 921-37. doi:10.1037/a0024700

Chua, J. L., Touyz, S., & Hill, a J. (2004). Negative mood-induced overeating in obese binge eaters: an experimental study. *International Journal of Obesity*, 28(4), 606–10. doi:10.1038/sj.ijo.0802595

Dittner, a J., Wessely, S. C., & Brown, R. G. (2004). The assessment of fatigue: a practical guide for clinicians and researchers. *Journal of psychosomatic research*, 56(2), 157–70. doi:10.1016/S0022-3999(03)00371-4

Eerola, T., Alluri, V., & Ferrer, R. (2008). Emotional Connotations of Isolated Instruments Sounds. *Energy*, (Icmpc 10), 483-489.

Eerola, T., Lartillot, O., & Toiviainen, P. (2009). Prediction of multidimensional emotional ratings in music from audio using multivariate regression models. In Proceedings of 10th International Conference on Music Information Retrieval (ISMIR 2009), Kobe, Japan, pp. 621–626.

Eich, E., Ng, J. T. W., Macaulay, D., Percy, A.D., & Grebneva, I. (2007). Combining

music with thought to change mood. In J.A. Coan and J.J.B. Allen (Eds.), *Handbook of emotion elicitation and assessment.* Series in affective science. (pp. 124-136). New York: Oxford University Press.

Ekman, P.(1992). Are there basic emotions. *Psychological Review.* 99: 550–553. doi: 10.1037/0033-295X.99.3.550

Ekman, P.(1992). "Basic Emotions" in *Handbook of Cognition and Emotion,* eds. T. Dalgleish and M. J. Power (pp. 45–60).. New York, NY:John Wiley and Sons.

Evans, P., & Schubert, E. (2008). Relationships between expressed and felt emotions in music. *Musicae Scientiae*, *12*(1), 75–99. doi:10.1177/102986490801200105

Feldman Barrett, L., & Russell, J. A. . (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology*, *74*(4), 967–984. doi:10.1037//0022-3514.74.4.967

Field, A. (2009). *Discovering statistics using SPSS, (3rd ed.)*. Sage publications.

Forgas, J. P., & George, J. M. (2001). Affective Influences on Judgments and Behavior in Organizations: An Information Processing Perspective. *Organizational Behavior and Human Decision Processes*, *86*(1), 3–34. doi:10.1006/obhd.2001.2971

Forgas, J. P. (1995). Mood and Judgment: The Affect Infusion Model (AIM). *Psychological Bulletin*, *117*(1), 39-66.

Gastwirth, J. L., Gel, Y. R., & Miao, W. (2009). The Impact of Levene's Test of Equality of Variances on Statistical Theory and Practice. *Statistical Science*, 24(3), 343–360. doi:10.1214/09-STS301

Gray, E. K., Watson, D. (2007). Assessing Positive and Negative Affect via Self-Report. In J.A. Coan and J.J.B. Allen (Eds.), *Handbook of emotion elicitation and assessment.* Series in affective science. (pp. 171-183). New York: Oxford University Press.

Gerrards-Hesse, A., Spies, K., Hesse, F. W. (1994). Experimental Inductions of Emotional States and their Effectiveness: A Review. *British Journal of Psychology*, *85*, 55-78.

Gomez, P., & Danuser, B. (2007). Relationships between musical structure and psychophysiological measures of emotion. *Emotion,* 7(2), 377–87. doi:10.1037/1528-3542.7.2.377

Hailstone, J. C., Omar, R., Henley, S. M. D., Frost, C., Kenward, M. G., & Warren, J. D. (2009). It's not what you play, it's how you play it: timbre affects perception of emotion in music. *Quarterly journal of experimental psychology (2006)*, *62*(11), 2141-55. doi:10.1080/17470210902765957

Henry, J. D., & Crawford, J. R. (2005). The short-form version of the Depression Anxiety Stress Scales (DASS-21): construct validity and normative data in a large non-clinical sample. *The British journal of clinical psychology* 44, 227–39. doi:10.1348/014466505X29657

Hevner, K. (1937). The affective value of pitch and tempo in music. *The American Journal of Psychology*, *49*(4), 621-630. University of Illinois Press.

Hu, X., Downie, J. S., Laurier, C., Bay, M., & Ehmann, A.F. (2008). The 2007 mirex audio mood classification task: Lessons learned. In Proceedings of the International Symposium on Music Information Retrieval (ISMIR), Philadelphia, PA, pp. 462–467.

Husain, G., Thompson, W. F., & Schellenberg, E. G. (2002). Effects of Musical Tempo and Mode on Arousal, Mood, and Spatial Abilities. *Music Perception*, *20*(2), 151-171. Univ California Press. doi:10.1525/mp.2002.20.2.151

Ilie, G.; Thompson, W. F. (2006). A Comparison of Acoustic Cues in Music and Speech for three Dimensions of Affect. *Music Perception*, 23(4), 319–330.

Juslin, P. N., Sloboda, J. (2010). Handbook of Music and Emotion: Theory, Research, Applications. USA: Oxford University Press. doi:10.1093/acprof:oso/9780199230143.001.0001

Johns, M. W. (1991). A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *sleep*, *14*(6), 540-545.

Johns, M. W. (1992). Reliability and factor analysis of the Epworth Sleepiness Scale. Sleep, 15, 376-381.

Kamenetsky, S. B., Hill, D. S., & Trehub, S. E. (1997). Effect of Tempo and Dynamics on the Perception of Emotion in Music. *Psychology of Music*, 25(2), 149–160. doi:10.1177/0305735697252005

Kallinen, K., & Ravaja, N. (2006). Emotion perceived and emotion felt: Same and different. *Musicae Scientiae*, *10*(2), 191–213. doi:10.1177/102986490601000203

Kawahara, J. I., & Sato, H. (2013). The effect of fatigue on the attentional blink. *Attention, perception & psychophysics*, *75*(6), 1096–102. doi:10.3758/s13414-013-0507-0

Kenealy, P. (1988). Validation of a Music Mood Induction Procedure : Some Preliminary Findings. *Cognition & Emotion*, *2*(1), 41-48. *Psychology Press*. doi:10.1080/02699938808415228

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: a review. *Biological psychology*, *84*(3), 394-421. doi:10.1016/j.biopsycho.2010.03.010

Krupp, L. B., LaRocca, N. G., Muir-Nash, J., & Steinberg, A. D. (1989). The fatigue severity scale: application to patients with multiple sclerosis and systemic lupus erythematosus. *Archives of neurology*, *46*(10), 1121.

Lang P. J., Bradley M. M., Cuthbert B. N. (1997). *International Affective Picture System (IAPS): Technical Manual and Affective Ratings.* NIMH Centre for the Study of Emotion and Attention, University of Florida.

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: affective, facial, visceral, and behavioral reactions. *Psychophysiology*, *30*(3), 261-73.

Lartillot, O. (2013). *MIRtoobox 1.5 User's Manual*. Retrieved from https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox

Lartillot, O. Eerola, T. Toiviainen, P. Fornari, J. (2008, September). *Multi-feature modeling of pulse clarity: Design, validation, and optimization*. International Conference on Music Information Retrieval, Philadelphia, PA.

Li, T., & Ogihara, O. M. (2003). Detecting emotion in music. *Proceedings of the International Symposium on Music Information Retrieval*, *John Hopkins University, Baltimore, MD,* 239–240.

Lichstein, K. L., Means, M. K., Noe, S. L., Aguillard, R. N. (1997). Fatigue and Sleep Disorders. *Behaviour Research and Therapy*, 35(8), 733–740.

Lieberman, H. R., Bathalon, G. P., Falco, C. M., Kramer, F. M., Morgan, C. a, & Niro, P. (2005). Severe decrements in cognition function and mood induced by sleep loss, heat, dehydration, and undernutrition during simulated combat. *Biological psychiatry*, *57*(4), 422–9. doi:10.1016/j.biopsych.2004.11.014

Lovibond, S. H. & Lovibond, P. F. (1995). *Manual for the Depression Anxiety Stress Scales. (2nd. Ed.)* Sydney: Psychology Foundation.

Lundqvist, L. O., Carlsson, F., Hilmersson, P., & Juslin, P. N. (2008). Emotional responses to music: experience, expression, and physiology. *Psychology of Music*, 37(1), 61–90. doi:10.1177/0305735607086048

Martin, M. (1990). On the Induction of Mood. *Clinical Psychology Review*, *10*, 669-697.

Montgomery, G. K. (1983). Uncommon tiredness among college undergraduates. *Journal of Consulting and Clinical Psychology*, *51*(4), 517-525.

Mayer, J. D., Allen, J. P., Beauregard, K. (1995). Mood Inductions for Four Specific Moods: A Procedure Employing Guided Imagery Vignettes With Music. *Journal of Mental Imagery, 199*(1&2), 133-150.

Montgomery, G. K. (1983). Uncommon tiredness among college undergraduates. *Journal of Consulting and Clinical Psychology*, *51*(4), 517–525.

Moore, B.C.J. (2010). Hearing. USA: Academic Press.

Niepel, G., Bibani, R. H., Vilisaar, J., Langley, R. W., Bradshaw, C. M., Szabadi, E., & Constantinescu, C. S. (2013). Association of a deficit of arousal with fatigue in multiple sclerosis: effect of modafinil. *Neuropharmacology*, *64*, 380–8. doi:10.1016/j.neuropharm.2012.06.036

Pagani, M., & Lucini, D. (1999). Chronic fatigue syndrome: a hypothesis focusing on the autonomic nervous system. *Clinical science*, *96*(1), 117–25.

Peretz, I., Gagnon, L., & Bouchard, B. (1998). Music and emotion: Perceptual determinants, immediacy, and isolation after brain damage. *Cognition, 68(2)*, 111–141.

Pignatiello, M. F., Camp, C. J., & Rasar, L. A. (1986). Musical mood induction: an alternative to the Velten technique. *Journal of abnormal psychology*, *95*(3), 295–7.

Posner, J., Russell, J. a, & Peterson, B. S. (2005). The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and psychopathology*, *17*(3), 715–34. doi:10.1017/S0954579405050340

Russo, F. A., Vempala, N. N., & Sandstrom, G. M. (2013). Predicting musically induced emotions from physiological inputs: linear and neural network models. *Frontiers in psychology*, *4*(August), 468. doi:10.3389/fpsyg.2013.00468

Russell J.A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology. 39*, 1161–1178. doi: 10.1037/h0077714

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*(1), 145-172. doi:10.1037/0033-295X.110.1.145

Sarason, I. G., Sarason, B. R., & Pierce, G. R. (1990). Anxiety, cognitive interference, and performance. *Journal of Social Behavior and Personality*, *5*(2), 1-18.

Sandstrom, G. M., & Russo, F. A. (2010). Music Hath Charms : The Effects of Valence and Arousal on Recovery Following an Acute Stressor. *Music and Medicine*, (2010), *2*(3), 137-143. doi:10.1177/1943862110371486

Sandstrom, G. M., & Russo, F. A. (2011). Absorption in music: Development of a scale to identify individuals with strong emotional responses to music. *Psychology of Music*, *41*(2), 216–228. doi:10.1177/0305735611422508

Schellenberg, E. G., Krysciak, A. M., & Campbell, R. J. (2000). Perceiving Emotion in Melody : Interactive Effects of Pitch and Rhythm. *Music Perception*, *18*(2), 155–171.

Schubert, E. (2004). Modeling perceived emotion with continuous musical features. *Music Perception*, *21*(4), 561–585.

Shahid, A., Shapiro, C. M., Wilkinson, K., & Marcu, S. (Eds.). (2012). *Stop, that and one hundred other sleep scales*. Springer.

Sharpe, M; Chalder, T., Palmer, I., Wessely, S. (1997). Chronic Fatigue Syndrome A Practical Guide to Assessment and Management. *General Hospital Psychiatry*, *19*, 185–199.

Sharpe, M., & Wilks, D. (2002). Fatigue. *BMJ (Clinical research ed.)*, *325*(7362), 480–3.

Shen, J., Barbera, J., & Shapiro, C. M. (2006). Distinguishing sleepiness and fatigue: focus on definition and measurement. *Sleep medicine reviews*, *10*(1), 63–76. doi:10.1016/j.smrv.2005.05.004

Skowronek, J., McKinney, M., & Par, S. (2007). A demostrator for automatic music mood estimation. Proceedings of the International Conference on Music Information Retrieval, Austria, 345–346.

Tellegen, A. and Atkinson, G. (1974) Openness to Absorbing and Self-altering Experiences ('Absorption'), a Trait Related to Hypnotic Susceptibility. *Journal of Abnormal Psychology, 83*, 268–77.

VanOyen Witvliet, C.; Vrana, S. R. (1995). Psychophysiological responses as indices of affective dimensions. *Psychophysiology*, *32*, 436-443.

Västfjäll, D. (2002). Emotion induction through music: A review of the musical mood induction procedure. *Musicae Scientiae (Special Issue),* 171-211.

Vuoskoski, J. K., & Eerola, T. (2011). Measuring music-induced emotion: A comparison of emotion models, personality biases, and intensity of experiences. *Musicae Scientiae*, *15*(2), 159-173. doi:10.1177/102986491140336.

Vytal, K. E., Cornwell, B. R., Letkiewicz, A. M., Arkin, N. E., & Grillon, C. (2013). The complex interaction between anxiety and cognition: insight from spatial and verbal working memory. *Frontiers in human neuroscience*, *7*(March), 93. doi:10.3389/fnhum.2013.00093

Watson, D (2000). *Mood and temperament*. New York: Guilford Press.

Witvliet, C. V. O., & Vrana, S. R. (2007). Play it again Sam: Repeated exposure to emotionally evocative music polarises liking and smiling responses, and influences other affective reports, facial EMG, and heart rate. *Cognition & Emotion*, *21*(1), 3–25. doi:10.1080/02699930601000672

Yang, Y., Lin, Y., Su, Y., & Chen, H. (2007). Music emotion classification: A regression approach. *In Proceedings of the IEEE International Confererence on Multimedia, Beijing, China*, 208–211.

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: characterization, classification, and measurement. *Emotion, 8*(4), 494–521. doi:10.1037/1528-3542.8.4.494

Zimny, G.H.; Weidenfeller, E.W. (1963). Effects of Music upon GSR and Heart-Rate. *The American Journal of Psychology*. *76*(2), 311-314.

Table 1

Descriptive statistics for age, and gender frequencies within each mood condition.

	Sadness		**Fatigue**		**Anxiety**	
Age	Mean	SD	Mean	SD	Mean	SD
	22.03	6	20.64	4.9	20.4	2.36
Gender	Male	Female	Male	Female	Male	Female
	5	32	14	27	9	28

Table 2

Number of participants who listened to each excerpt by condition.

Condition	Track	N
Anxiety	S44	6
	S47	5
	S50	7
	S23	9
	S27	7
	S34	6
Fatigue	S11	8
	S19	8
	S35	8
	S37	8
	S49	6
	SP7	8
Sadness	S13	6
	S16	7
	S25	8
	S39	6
	S40	7
	S52	7

Table 3

List of tracks used in Mood Induction Phase with subjective ratings post induction

Condition	Study Track Number	Track Title	Album	Label	Time code	Valence		Arousal		Tired		Sad		Anxious	
						Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Anxiety	23	The Day Comes	Battlestar Galactica	La-La Land Records	00:00-01:00	6.33	1.73	2.67	1.41	55.17	15.39	4.11	4.46	4.61	3.18
Anxiety	27	Stampede	Australia	Fox Music	00:00-01:00	5.29	1.98	4.71	3.2	54	35.99	11.43	7.23	30.07	27.85
Anxiety	34	Thousands Left Behind	Battlestar Galactica	La-La Land Records	00:00-01:00	4.83	1.47	4.17	2.22	60.08	33.53	27.5	16.47	51.92	17.17
Anxiety	44	Louis' Revenge	Interview with a Vampire	Geffen Records	00:00-01:00	5.17	1.6	4	1.55	50.5	35.05	11.33	16.31	28.83	28.88
Anxiety	50	Firenze di Notte	Hannibal OST	Decca	00:00-01:00	5.14	1.68	3	2.53	58.64	33.55	18.43	12.24	23.5	30.69
Anxious	47	Devil's Monsters	Halo – Combat Evolved OST	Sumthing Else Music Works, Inc	00:00-01:00	4	2	4.5	1.73	65	9.56	24.13	22.54	51.87	21.22
Fatigue	11	Reunion	Coraline OST	Koch Records	00:00-01:00	5.37	1.6	2.88	1.36	66.38	30.17	25.94	25.21	19.56	15.12
Fatigue	19	Earth	Gladiator OST	Decca	00:00-01:00	6.5	2.62	2.13	1.64	49	32.2	17.69	19.06	20.25	26.55
Fatigue	35	Mary's Museum	Equilibrium OST	Unofficial Release	00:00-01:00	4.5	1.07	3	1.41	59.13	27.5	31.88	17.64	41.44	20.94
Fatigue	37	Good Night	Battlestar Galactica	La-La Land Records	00:00-01:00	5.58	1.27	2.57	1.51	69.5	27.88	29	24.54	30.38	22.95
Fatigue	49	Is Possible	Equilibrium OST	Unofficial Release	00:00-01:00	5.8	1.1	2.6	1.14	51.75	42.39	10.57	12.03	17.42	13.55
Fatigue	SP7	Saying Goodbye to Those You Love	A Beautiful Mind	Decca	00:38-01:38	5.63	2.13	2.75	1.75	57.69	36.36	22	20.2	18.13	24
Sadness	13	Silk Road	Crouching Tiger, Hidden Dragon	Sony	00:00-01:00	5.17	0.98	4.33	1.97	67.66	20.12	19.75	17.95	35	26.59
Sadness	16	Virtue	Hannibal OST	Decca	00:00-01:00	6	1.26	3.67	1.5	48.79	29.57	5.36	3.29	16.5	16.12
Sadness	25	Election by Adoration	Angels and Demons	Sony Classics	00:00-01:00	5.63	3.29	2.5	1.93	60.88	33.51	29	23.63	42.75	30.79
Sadness	39	A Gift of Thistle	Braveheart	Decca	00:00-01:00	5	1.26	2.17	2.04	53.17	19.72	29.91	30.59	27.25	26.52
Sadness	40	A Small Measure of Peace	The Last Samurai OST	Elektra	00:00-01:00	5.14	2.34	2.71	2.29	53.17	32.37	47.67	33.21	43.5	32.83
Sadness	52	Murion's Burial	Braveheart	Decca	00:00-01:00	6.14	0.9	3.29	1.71	40.5	40.37	10.64	12.24	12.57	12.87

Table 4

List of tracks used in the follow up pilot phase

Title	Album	Track Number	Label	Timecode	Tired	Sad	Anxious
Haunting & Heartbreaking	Lost Highway	5	Interscope	00:31:13-01:31:13	1.77	2.88	4.34
The Way of the Sword	The Last Samurai: Original Motion Picture Score	10	Elektra Records	06:20:13-07:20:13	2.67	3.36	1.09
Corynorhinus	Batman Begins: Original Motion Picture Soundtrack	11	Warner Bros. Records	00:00:00-01:00:00	3.13	4.41	1.4
The English Patient	The English Patient	1	Fantasy Records	00:00:00-01:00:00	2.9	2.6	1.6
A Tale (Un cuento)	Pan's Labyrinth	13	Milan Entertainment	00:53:00-01:53:00	2.04	3.69	2.69
Main Title Theme	Mullholand Drive: Original Motion Picture Score	2	Milan Entertainment	00:14:00-01:14:00	1.64	3.04	2.58
Saying Goodbye to Those You Love	A Beautiful Mind	12	Decca	00:38:15-01:38:15	4.83	2.29	0.39
To the Stars	DragonHeart: Original Motion Picture Soundtrack	2	MCA Records	00:00:00-01:00:00	3.4	2.9	2.9

Danilov's Confession	Enemy at the Gates	11	Sony Records	00:00:00-01:00:00	2.14	3.5	0.98
Gasoline	Pay it Forward	24	Varese Sarabande	00:16:21-01:16:21	3.48	3.12	0.61
Those We Don't Speak Of	The Village	4	Hollywood	00:39:01-01:39:01	0.569	1.38	4.46
Ghosts	Road to Perdition	24	Decca Records	00:45:00-01:45:00	4.06	3.47	0.78

Table 5

Signal Processing Settings.

Feature	High Pass Filter	Low Pass Filter	Other filters	Bin window	Calculation method
Heart rate	0.5 Hz	4 Hz	none	20 seconds	bin rates
Respiration rate	0.05Hz	1 Hz	none	20 seconds	bin rates
Skin conductance level	None	None	none	20 seconds	bin means
Corrugator	5 Hz	500 Hz	notch filter (59-61Hz), RMS50	20 seconds	bin means
Zygomaticus	5 Hz	500 Hz	notch filter (59-61Hz), RMS50	20 seconds	bin means

Table 6

Descriptives statistics

Condition	**DASS D**		**DASS A**		**DASS S**		**ESS**		**FSS**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Sadness	4.56	3.61	5.42	3.94	7.97	4.70	9.68	3.65	36.05	11.74
Fatigue	4.00	3.92	4.29	4.15	6.33	3.61	8.24	3.71	33.5	8.05
Anxiety	4.06	4.58	3.86	3.43	7.00	4.80	8.58	3.32	33.12	10.31

Note: The Depression, Anxiety and Stress Scale (DASS21) from Lovibond, & Lovibond (1995); The Epworth Sleep Sleepiness Scale (ESS) from Johns (1991); The Fatigue Severity Scale from Krupp et al. (1989).

Table 7

Mean valence and arousal ratings post induction, by track.

Condition	Track	Valence Mean	SD	Arousal Mean	SD
Sadness	S13	5.17	0.98	4.33	1.97
	S16	4.83	1.26	3.67	1.50
	S25	5.63	3.29	2.50	1.93
	S39	5.00	1.26	2.17	2.04
	S40	5.14	2.34	2.71	2.29
	S52	6.14	0.90	3.29	1.70
Fatigue	S11	5.38	1.60	2.88	1.36
	S19	6.50	2.62	2.13	1.64
	S35	4.50	1.07	3.00	1.41
	S37	5.57	1.27	2.57	1.51
	S49	5.80	1.09	2.60	1.14
	SP7	5.63	2.13	2.75	1.75
Anxiety	S23	6.33	1.73	2.67	1.41
	S27	5.29	1.98	4.71	3.20
	S34	4.83	1.47	4.17	2.23
	S44	5.17	1.60	4.00	1.55
	S47	4.00	2.00	4.50	1.73
	S50	5.14	1.68	3.00	2.53

Table 8

Summary of ANOVA results for all physiological measures in 4 minute moving widows

Measure	**1-4 min**	**3-6 min**	**5-8 min**	**7-10 min**	**9-12 min**	**11-14 min**	**13-16 min**	**15-18 min**	**17-20 min**	**20 min**
	F = 1.34, *p* .267	F = 1.48, *p* = .235,	F = 1.75, *p* = .182	F = 1.6, *p* = .209	F = 1.04, *p* = .358,	F = 1.2, *p* = .308,	F = 1.6, *p* = .209	F = 1.87, *p* = .162,	F = 1.77, *p* = .178	F = 1.6, *p* = .209
Heart rate	T<S=A	T<S<A	T<S<A	T<S<A	T<S<A	T<S<A	T<S=A	T<S=A	T<S=A	T<S<A
	F = .081,	F = .011,	F = .079,	F = .186 *p*	F = .139,	F = .353,	F = .47,	F = .29,	F = .694,	F = .151,
Respiration	*p* = .922,	*p* = .99	*p* = .924,	= .831,	*p* = .870,	*p* = .704,	*p* = .621,	*p* = .749,	*p* = .503	*p* = .86,
	T=S=A	T=S=A	T=S=A	T<S<A	T=A<S	A<T=S	A<S=T	T=S=A	T=A<S	T=S=A
Skin	F = 2.47,	**F = 3.44,**	**F = 4.28,**	**F = -4.13,**	F = 2.67,	F = .691,	F = .364,	F = 1.47,	F = 2.72,	F = 2.3,
Conductance	*p* = .092,	***p* = .038,**	***p* = .018,**	***p* = .02,**	*p* = .077	*p* = .504,	*p* = .69,	*p* = .237,	*p* = .073,	*p* = .109
Level	T<S=A	**T<A<S**	**T<A<S**	**T<A<S**	T<S=A	T<S=A	T<S=A	T<A<S	T<A<S	T<S=A
	F = 1.93,	F = 2.49,	F = 2.34,	F = 2.82,	F = 1.82,	F = ,762,	F = .482,	F = .658,	F = .809,	F = 1.68,
	p=.153,	*p*=.119,	*p*=.104,	*p*=.097,	*p*=.170	*p*=.471,	*p*=.620,	*p*=.52,	*p*=.45,	*p*=.193,
Corrugator	T<S<A	T<S=A	T<S=A	T<S=A	T<S=A	T<S=A	T=S=A	T<A<S	T<A<S	T<S=A
	F = 1.2,	F = .359,	F = .483,	F = .668,	F = .327,	F = .207,	F = .209,	F = .17,	F = .532,	F = .51,
	p =.30,	*p* = .7,	*p* = .619,	*p* =. 516,	*p* = .722,	*p* = .813,	*p* = .812,	*p* = .84,	*p* = .59,	*p* = .603,
Zygomatic	T<S<A	T<A<S	T=S=A	T=S=A	T=S=A	T<S=A	T=S=A	T=S=A	S<T<A	T<S<A

Table 9

Multiple Regression Coefficients for Questionnaire Scores Predicting Heart Rate.

	B	**SE B**	***B***
Step 1			
Constant	0.739	0.393	
DASSD	-0.018	0.038	-0.081
DASSA	0.011	0.03	0.049
DASSS	0.004	0.033	0.019
ESS	-0.012	0.027	-0.048
FSS	-0.018	0.011	-0.201
Step 2			
Constant	0.749	0.38	
DASSD	-0.016	0.031	-0.07
DASSA	0.012	0.028	0.055
ESS	-0.012	0.027	-0.049
FSS	-0.018	0.011	-0.201
Step 3			
Constant	0.776	0.374	
DASSD	-0.008	0.026	-0.037
ESS	-0.011	0.027	-0.045
FSS	-0.018	0.011	-0.204
Step 4			
Constant	0.785	0.371	
ESS	-0.011	0.027	-0.043
FSS	-0.02	0.01	-0.22
Step 5			
Constant	0.721	0.334	
FSS	-0.021	0.009	-0.23*

Note: $R^2 = .058$ for step 1, $\Delta R^2 < .001$ for step 2 (ns), $\Delta R^2 = -.002$ for step 3 (ns), $\Delta R^2 = -.001$ for step 4 (ns),), $\Delta R^2 = -.002$ for step 5 (ns). $^*p < .05$.

The Depression, Anxiety and Stress Scale (DASS21) from Lovibond, & Lovibond (1995); The Epworth Sleep Sleepiness Scale (ESS) from Johns (1991); The Fatigue Severity Scale from Krupp et al. (1989).

Table 10

Multiple Regression Coefficients for Questionnaire Scores Predicting Skin Conductance Level.

	B	**SE B**	***B***
Step 1			
Constant	-0.186	0.581	
DASSD	-0.055	0.056	-0.169
DASSA	0.023	0.045	0.072
DASSS	-0.036	0.049	-0.128
ESS	0.034	0.039	0.095
FSS	0.003	0.016	0.022
Step 2			
Constant	-0.111	0.414	
DASSD	-0.052	0.054	-0.16
DASSA	0.023	0.045	0.072
DASSS	-0.035	0.048	-0.127
ESS	0.035	0.038	0.099
Step 3			
Constant	-0.101	0.412	
DASSD	-0.046	0.052	-0.143
DASSS	-0.026	0.045	-0.095
ESS	0.036	0.037	0.102
Step 4			
Constant	-0.198	0.376	
DASSD	-0.07	0.034	-0.215
ESS	0.037	0.037	0.104
Step 5			
Constant	0.123	0.193	
DASSD	-0.067	0.034	-0.207*

Note: R^2 = .061 for step 1, ΔR^2 = .016 for step 2 (ns), ΔR^2 = .024 for step 3 (ns), ΔR^2 = -.032 for step 4 (ns),), ΔR^2 = -.032 for step 5 (ns). *p = .051.

The Depression, Anxiety and Stress Scale (DASS21) from Lovibond, & Lovibond (1995); The Epworth Sleep Sleepiness Scale (ESS) from Johns (1991); The Fatigue Severity Scale from Krupp et al. (1989).

Table 11

Multiple Regression Coefficients for Questionnaire Scores Predicting Zygomatic Activity.

	B	**SE B**	**B**
Step 1			
Constant	-0.024	0.023	
DASS-D	-0.004	0.002	-0.302*
DASS-A	-0.004	0.002	-0.282*
DASS-S	0.004	0.002	0.303
EPWORTH	0.001	0.002	0.078
FATIGUE	0	0.001	0.095
Step 2			
Constant	-0.017	0.021	
DASS-D	-0.004	0.002	-0.306*
DASS-A	-0.004	0.002	-0.279*
DASS-S	0.004	0.002	0.302
FATIGUE	0.001	0.001	0.112
Step 3			
Constant	0.001	0.01	
DASS-D	-0.003	0.002	-0.254
DASS-A	-0.004	0.002	-0.279*
DASS-S	0.003	0.002	0.295

Note: $R^2 = .114$ for step 1, $\Delta R^2 = -.006$ for step 2 (ns), $\Delta R^2 = -.010$ for step 3 (ns).

*p < .05

The Depression, Anxiety and Stress Scale (DASS21) from Lovibond, & Lovibond (1995); The Epworth Sleep Sleepiness Scale (ESS) from Johns (1991); The Fatigue Severity Scale from Krupp et al. (1989).

Table 12

Summary of significant Levene's test values for 4 minute SCL comparisons

Time point	**Levene's test F statistic**	**sig.**
T7	$F(2.67) = 4.876$	$p = .011$
T10	$F(2,67) = 3.273$	$p = .044$
T11	$F(2, 67) = 3.797$	$p = .027$
T15	$F(2,67) = 3.269$	$p = .044$
T16	$F(2, 67) = 3.730$	$p = .029$
T20	$F(2,67) = 3.595$	$p = .033$

Table 13

Summary and description of audio features extracted using MIRToolbox

Feature	Description
RMS (root mean square)	Level of energy within the signal. Computed by taking the the square root of the arithmetic mean of the squared voltage abstracted from the waveform.
Low energy	Indicator of the temporal distribution of energy in the signal. Computed using the percentage of frames showing less-than-average energy.
Event density	Average frequency of events, such as the number of note onsets per unit of time.
Tempo	Estimate of tempo. Computed by identifying periodicities in the signal.
Pulse clarity	Estimate of rhythmic clarity, strength of beats.
Zero cross	Measure of noise. Computed by counting the number of instances the signal crosses the x axis.
Centroid	The centroid of the spectral distribution.
Spread	The spectral spread of the sound signal.
Rolloff	Frequency cutoff below which 85% of the spectral energy in the signal is contained.
Brightness	Amount of spectral energy above 1500 Hz.
Irregularity	Amount of variation in the successive peaks of a spectrum.
Mode	Estimate of major vs. minor mode in the form of a value between -1 and +1, where more negative values indicate a major, and more positive values indicate a minor mode.

Table 14

Multiple regression coefficients for audio features predicting arousal

	B	SE B	*B*
Step 1			
Constant	0.697	1.114	
RMS	-8.012	10.181	-0.148
Pulse Clarity	0.748	1.096	0.089
Spread	-0.956	0.438	-0.375*
Step 2			
Constant	0.698	1.11	
RMS	-5.151	9.244	-0.095
Spread	-0.931	0.435	-0.365*
Step 3			
Constant	0.142	0.482	
Spread	-0.747	0.281	-0.293*

Note: R^2 = .095 for step 1, ΔR^2 = -.006 for step 2 (ns), ΔR^2 = -.004 for step 3 (ns).

*p < .05

Table 15

Multiple regression coefficients for audio features predicting anxiety

	B	SE B	*B*
Step 1			
Constant	18.189	12.248	
RMS	-29.534	111.934	-0.048
Pulse Clarity	-10.829	12.046	-0.113
Spread	-10.842	4.816	-0.375*
Step 2			
Constant	15.54	6.978	
Pulse Clarity	-12.138	10.912	-0.127
Spread	-9.921	3.299	-0.343*
Step 3			
Constant	10.501	5.316	
Spread	-8.672	3.106	-0.3*

Note: R^2 = .105 for step 1, ΔR^2 = -.001 for step 2 (ns), ΔR^2 = -.014 for step 3 (ns).

*p < .05

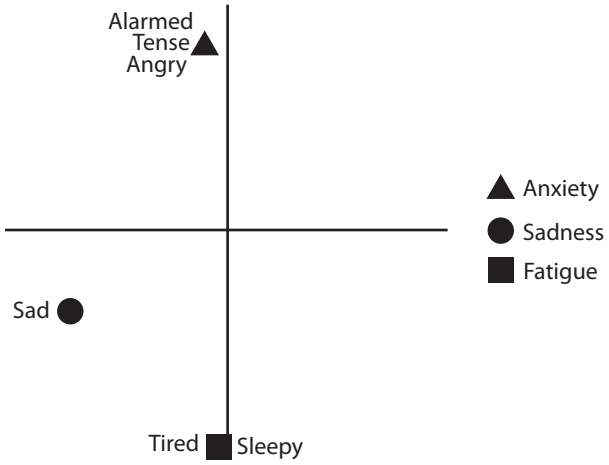

Figure 1. Russell's circumplex of affect (Adapted from Russell, 1980, p. 1167).

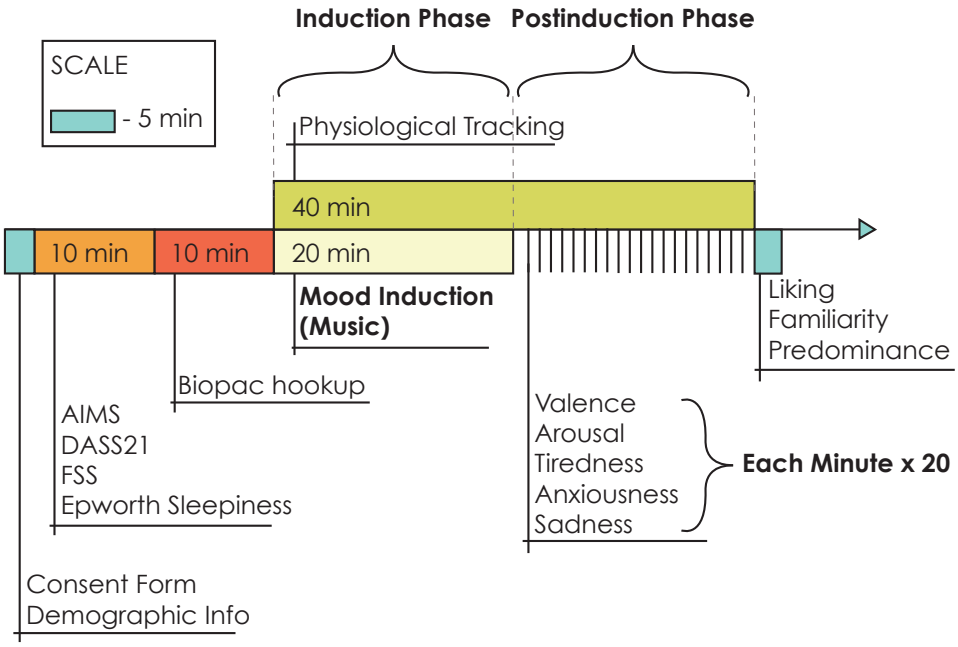

Figure 2. Graphic representation of the mood induction procedure.

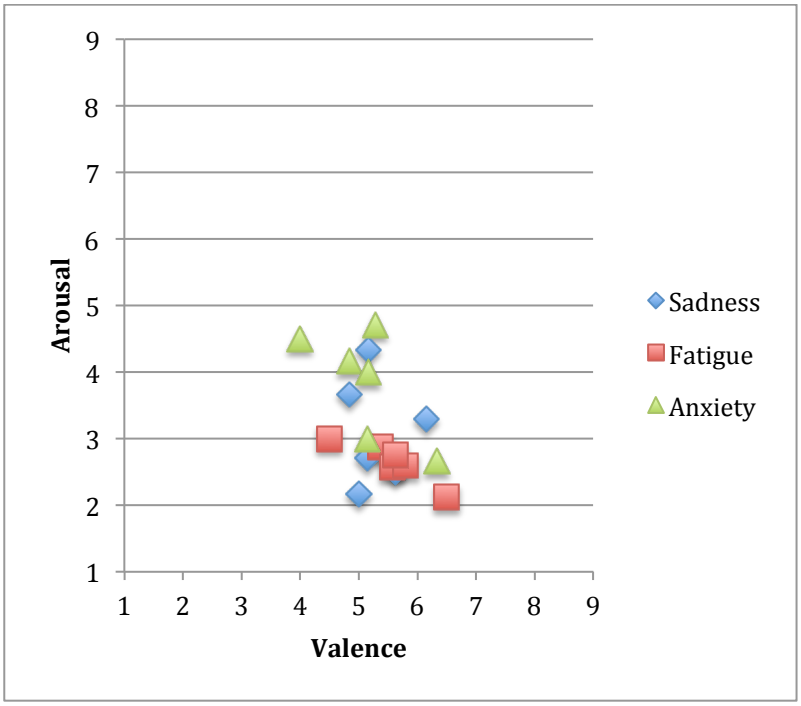

Figure 3. Excerpts (by condition) within the valence/arousal space based on participant ratings post induction.

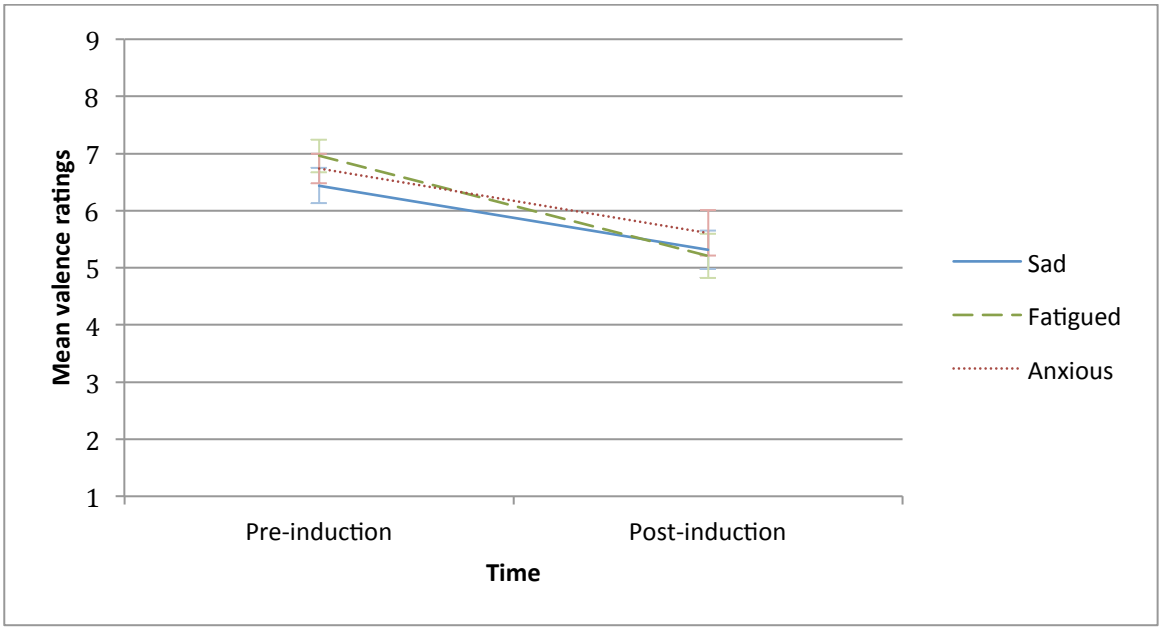

Figure 4. Valence ratings pre and post mood induction. All conditions showed a significant reduction in valence ratings after music listening.

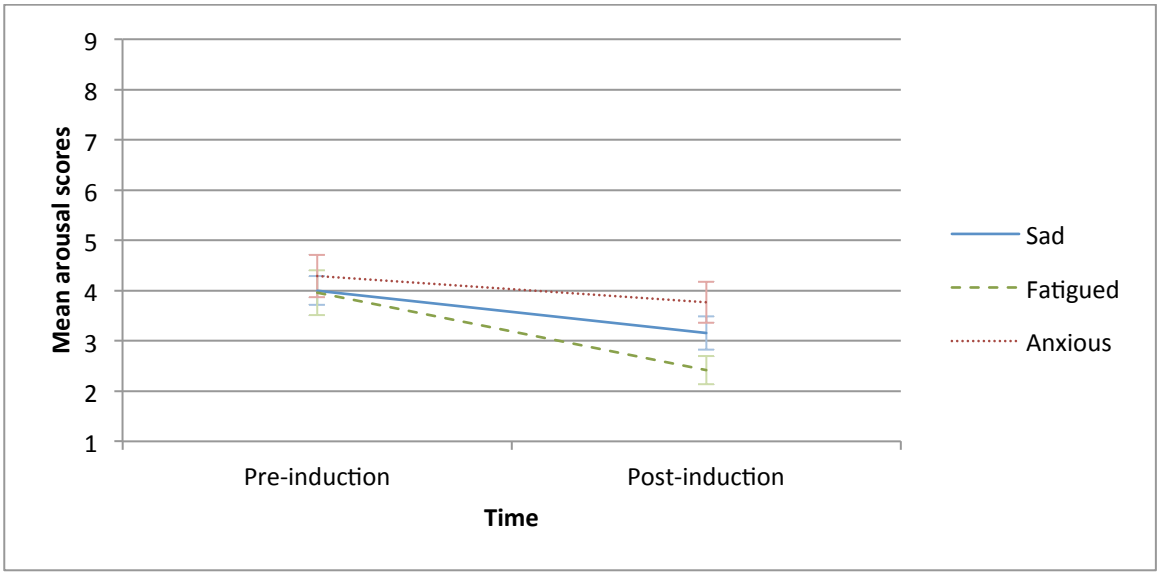

Figure 5. Arousal ratings pre and post mood induction. All conditions showed a significant reduction in arousal after music listening.

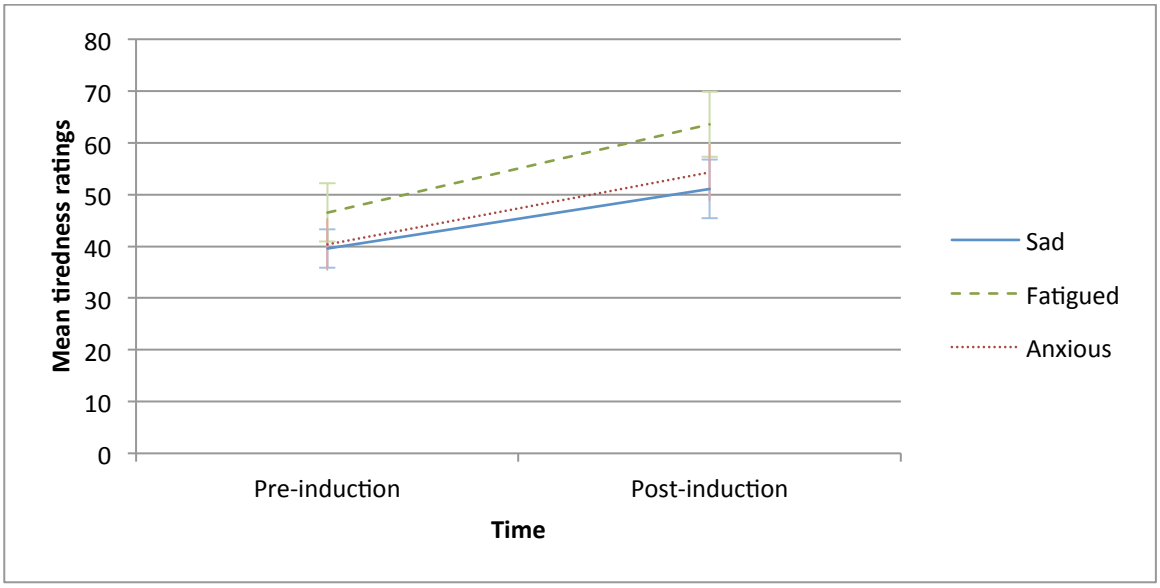

Figure 6. Tiredness ratings pre and post mood induction. All conditions showed an increase in tiredness after music listening.

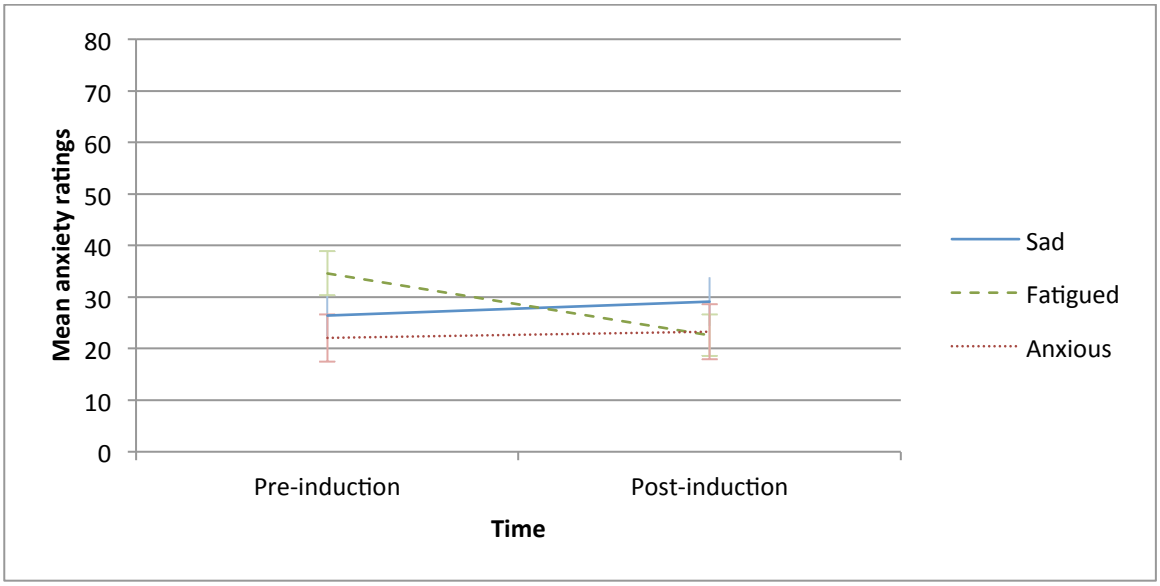

Figure 7. Anxiety ratings pre and post mood induction. The sadness and anxiety conditions showed a slight increase in anxiety ratings, while the fatigue condition showed a decrease.

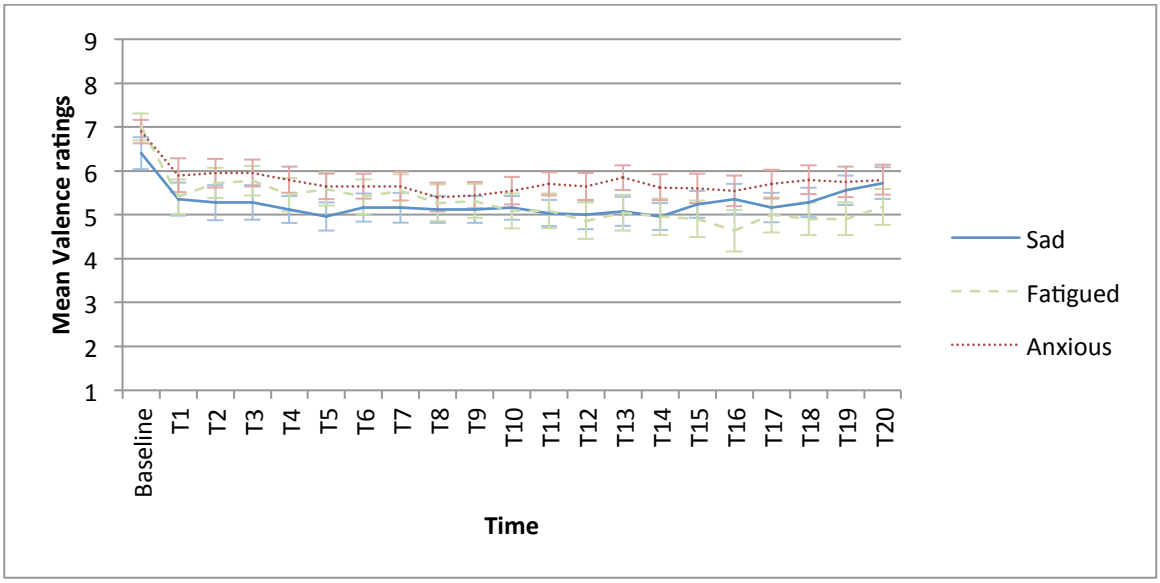

Figure 8. Valence ratings over time. Induction phase occurs between baseline and T1, postinduction phase between T1 and T20. After the initial reduction in valence in all conditions following the mood induction period, ratings remained relatively stable over the 20-minute postinduction period.

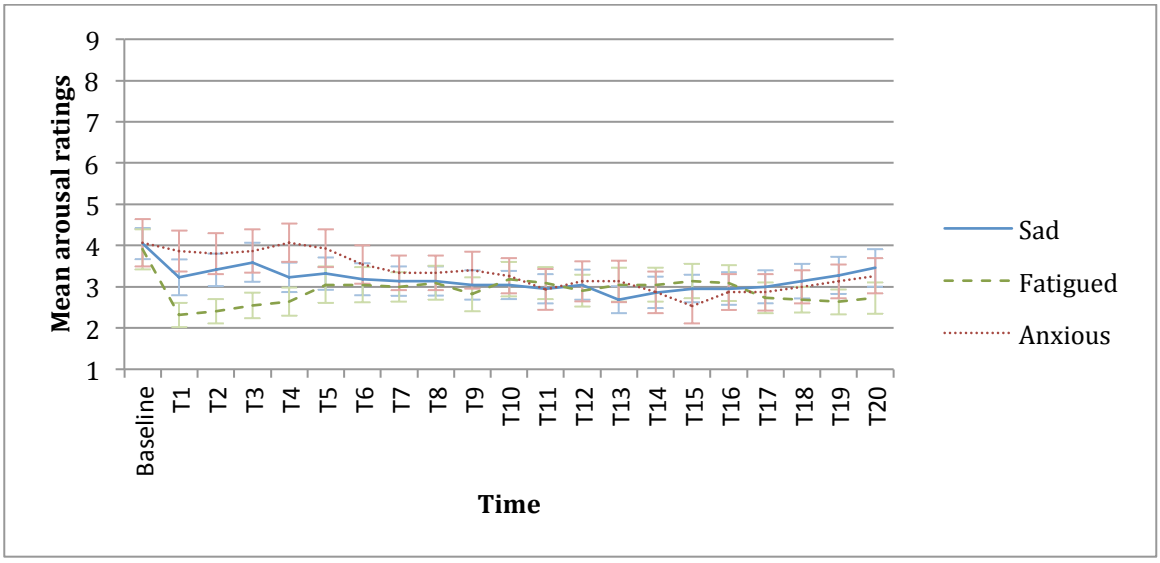

Figure 9. Arousal ratings over time. Induction phase occurs between baseline and T1, postinduction phase between T1 and T20. Following the initial 5 minutes of the postinduction period, participants reported similar levels of arousal across groups, which were lower than at baseline.

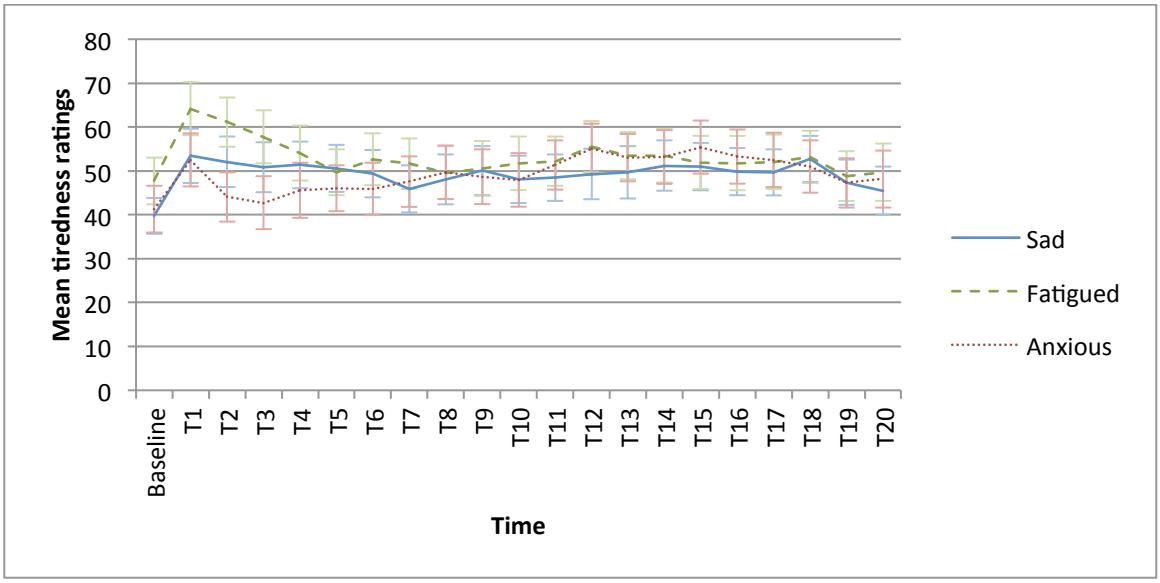

Figure 10. Tiredness ratings over time. Induction phase occurs between baseline and T1, postinduction phase between T1 and T20. All groups showed an increase in tiredness ratings, with the most pronounced change evident in the fatigue condition. After a period of roughly 5 minutes, all groups began to exhibit similar tiredness ratings.

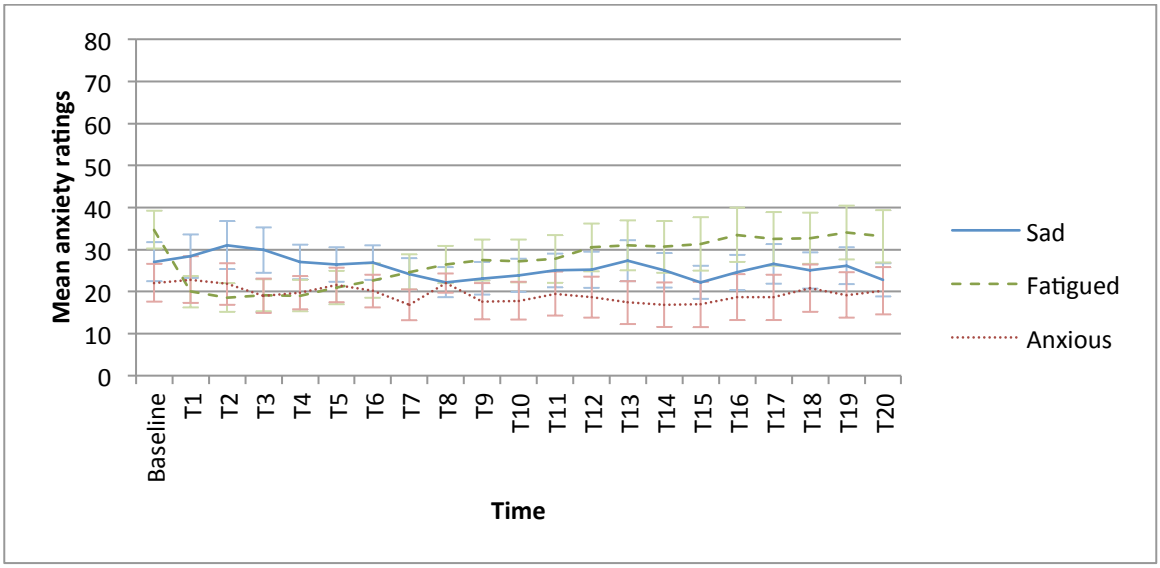

Figure 11. Anxiety ratings over time. Induction phase occurs between baseline and T1, postinduction phase between T1 and T20. In contrast with the sadness and anxiety conditions, which showed little fluctuation in ratings over time, the fatigue condition exhibited an initial decrease in anxiety ratings that was sustained for roughly 5 minutes, following which ratings gradually returned to baseline.

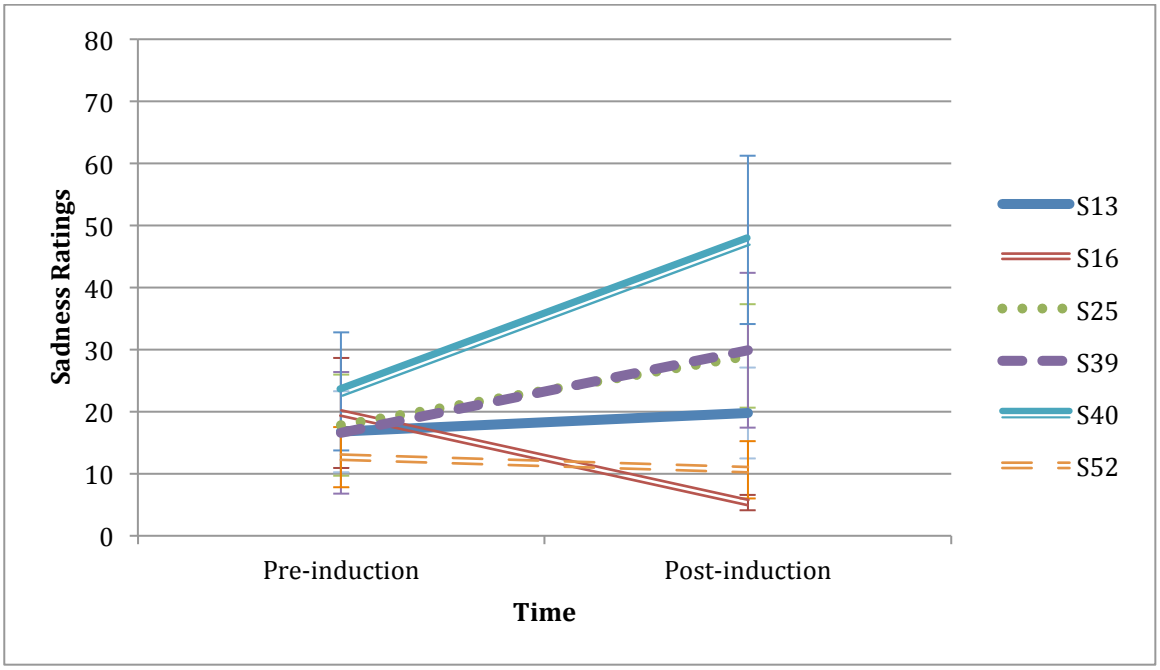

Figure 12. Sadness ratings by track in the sadness condition.

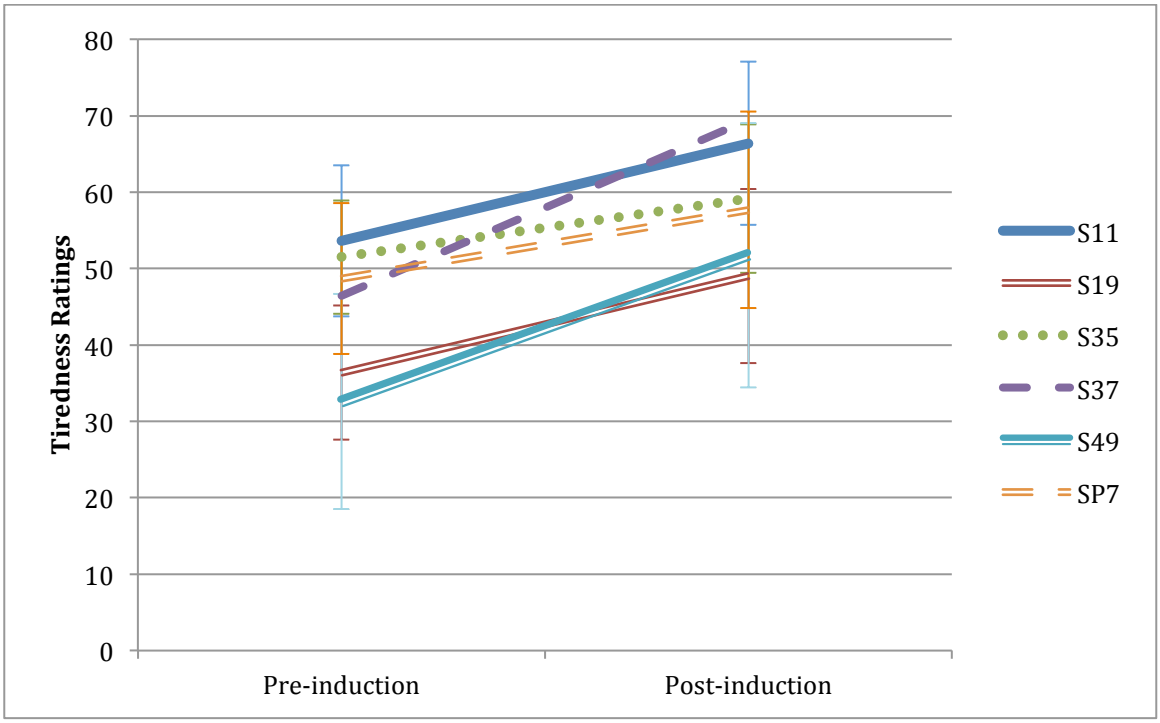

Figure 13. Tiredness ratings by track in the fatigue condition.

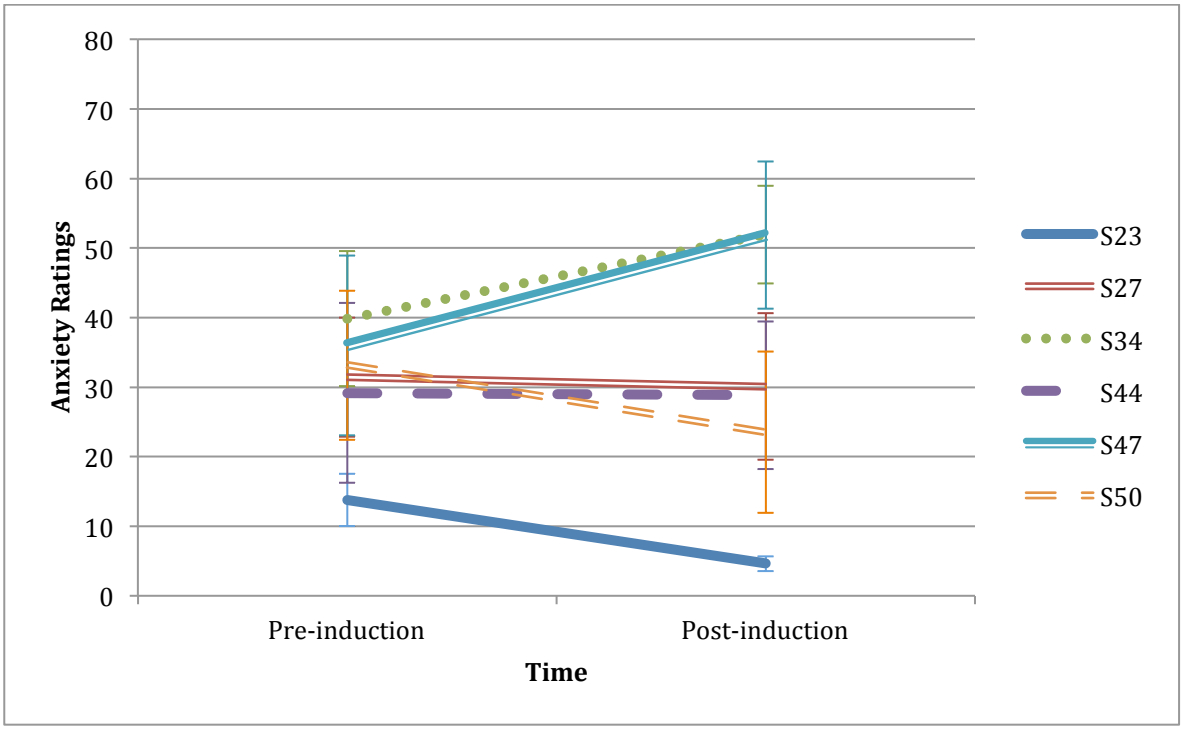

Figure 14. Anxiety ratings by track in the anxiety condition.

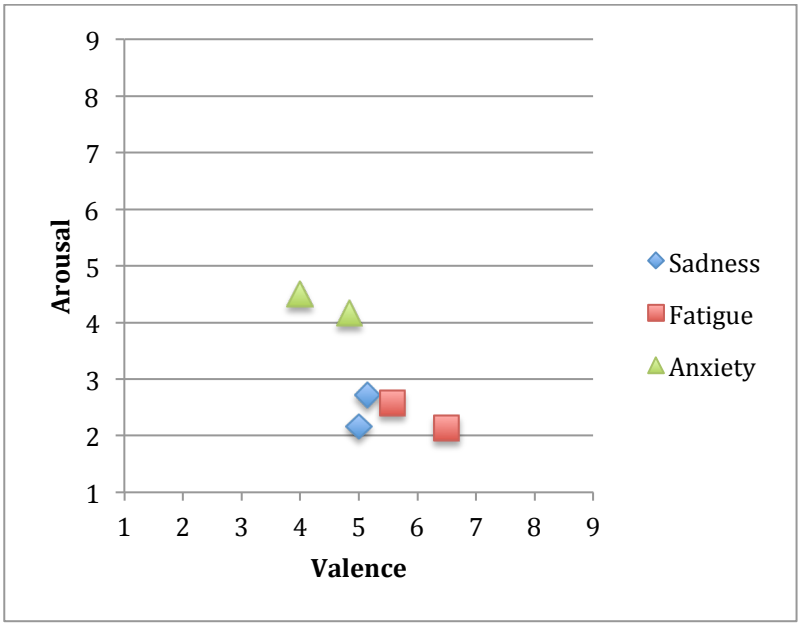

Figure 15. The two tracks most successful at eliciting intended mood based on target mood ratings, depicted within the valence/arousal space based on corresponding ratings of valence and arousal post induction.

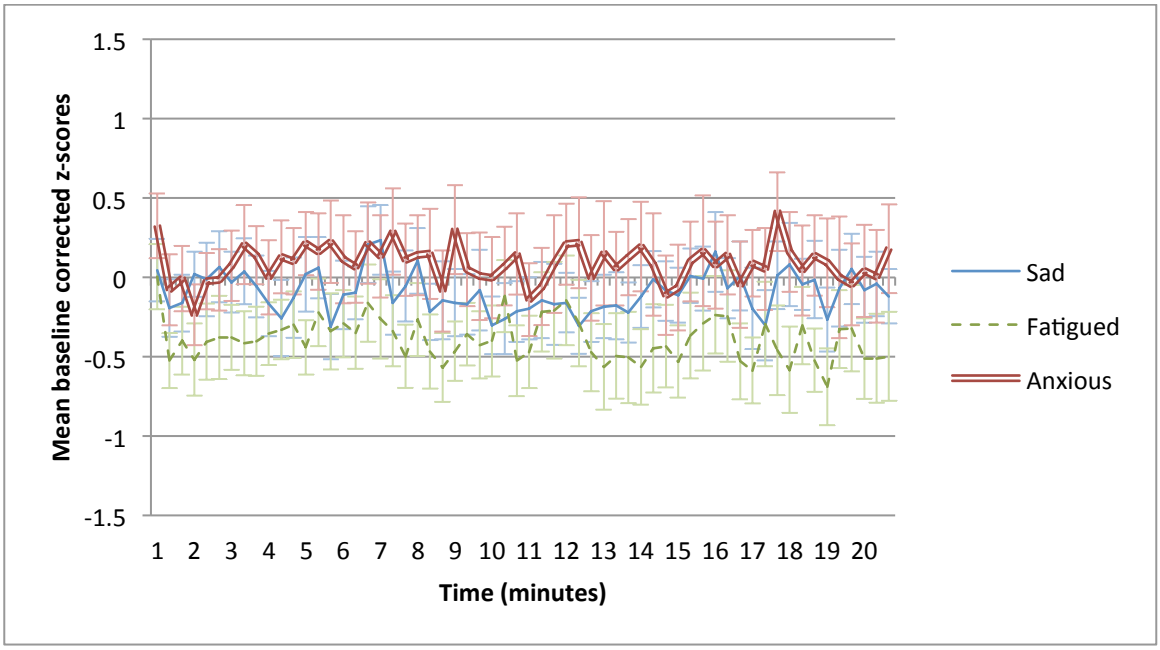

Figure 16. Heart rate data (represented in baseline corrected z scores) with standard errors over the course of the 20 minute mood induction period.

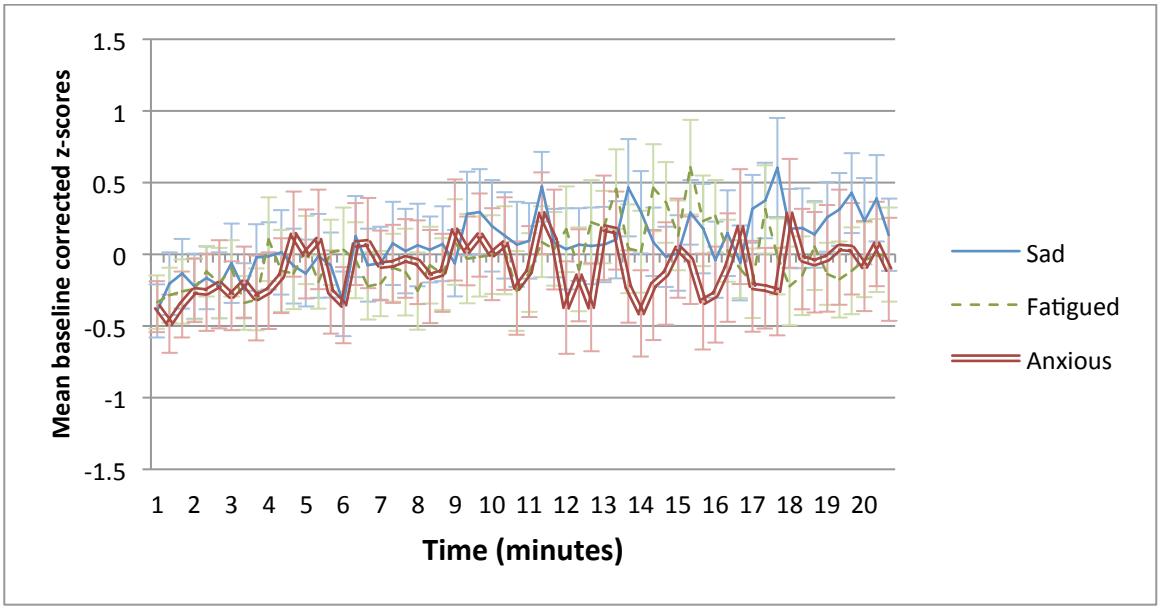

Figure 17. Respiration rate data (represented in baseline corrected z scores) over the course of the 20 minute mood induction period.

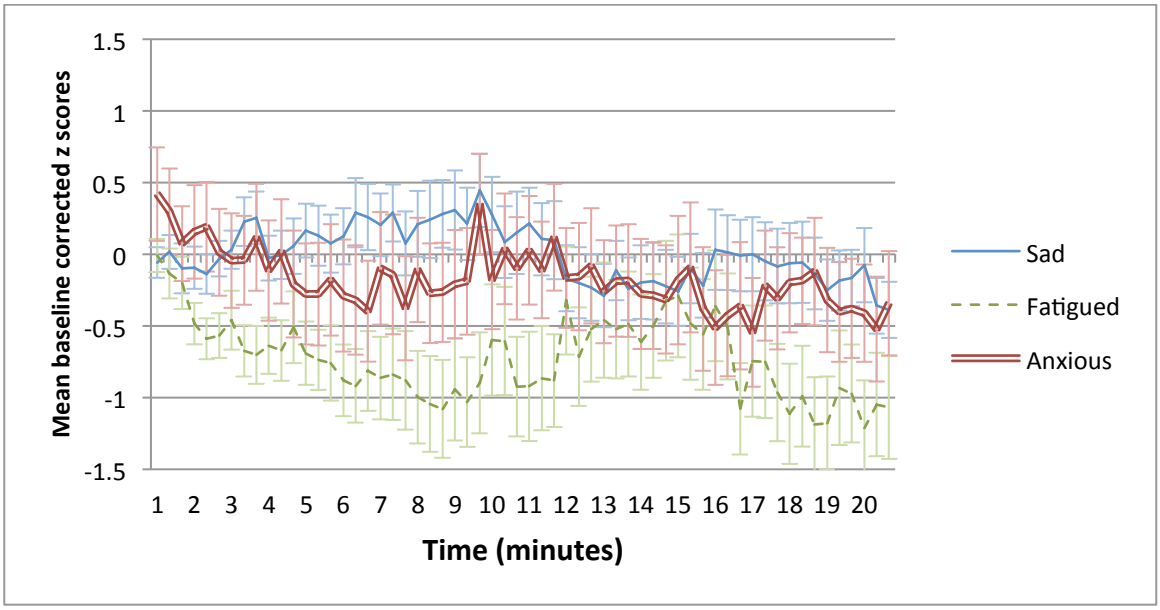

Figure 18. Skin conductance level data (represented in baseline corrected z scores) over the course of the 20 minute mood induction period. Overall SCL was lowest in the fatigue condition. The sadness and anxiety conditions showed similar levels of SCL, however, while the anxiety condition exhibited a decrease in SCL, the sadness condition exhibited an increase over the first 7 or so minutes.

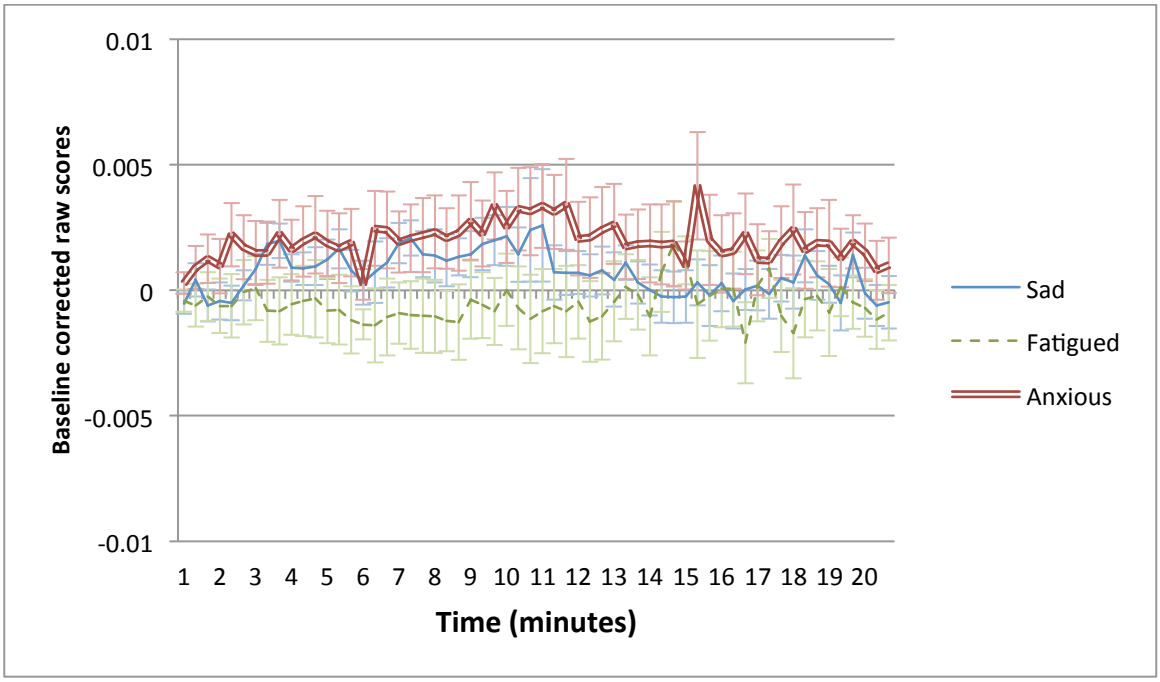

Figure 19. Corrugator data (baseline corrected raw scores) over the course of the 20 minute mood induction period.

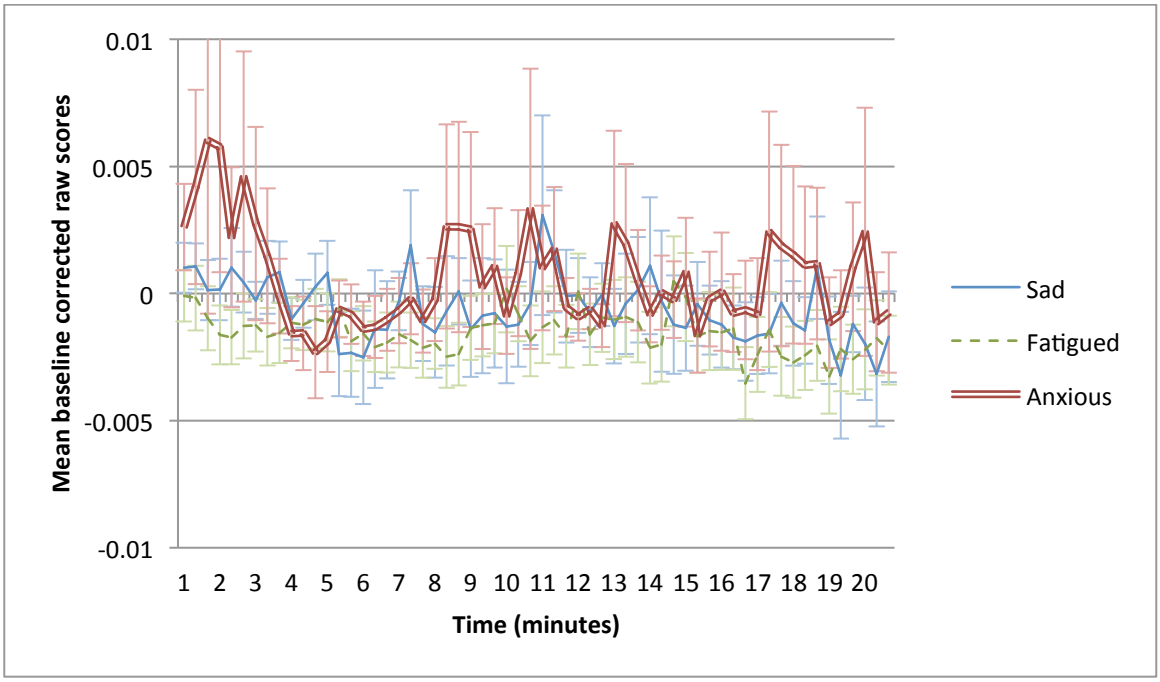

Figure 20. Zygomaticus data (baseline corrected raw scores) over the course of the 20 minutes mood induction period. Groups seem to exhibit similar levels of zygomaticus activity over time, although the fatigue condition seems to show slightly less activity overall.