Modulation of the startle reflex by pleasant and unpleasant music

Mathieu Roy, Jean-Philippe Mailhot, Nathalie Gosselin, Sébastien Paquette, Isabelle Peretz*

Department of Psychology, BRAMS, University of Montreal, Canada

ARTICLE INFO
Available online 23 July 2008

Keywords:
Music
Emotion
Startle reflex
Zygomatic
Corrugator
Skin conductance
Heart rate

ABSTRACT

The issue of emotional feelings to music is the object of a classic debate in music psychology. Emotivists argue that emotions are really felt in response to music, whereas cognitivists believe that music is only representative of emotions. Psychophysiological recordings of emotional feelings to music might help to resolve the debate, but past studies have failed to show clear and consistent differences between musical excerpts of different emotional valence. Here, we compared the effects of pleasant and unpleasant musical excerpts on the startle eye blink reflex and associated body markers (such as the corrugator and zygomatic activity, skin conductance level and heart rate). The startle eye blink amplitude was larger and its latency was shorter during unpleasant compared with pleasant music, suggesting that the defensive emotional system was indeed modulated by music. Corrugator activity was also enhanced during unpleasant music, whereas skin conductance level was higher for pleasant excerpts. The startle reflex was the response that contributed the most in distinguishing pleasant and unpleasant music. Taken together, these results provide strong evidence that emotions were felt in response to music, supporting the emotivist stance.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

The emotional power of music remains a mystery. Unlike most emotional inducers, music is not a sentient being nor does it seem to have any obvious adaptive value (Pinker, 1997). Yet, most people affirm that they feel strong emotions when they listen to music (Sloboda and O'Neill, 2001). This paradox led many music scholars to believe that music is only iconic or representative of emotion, a position coined as ‘cognitivist’ by Kivy (1990). Opponents to this view, known as ‘emotivists’, feel that the cognitivist position does not render justice to the direct and unmediated fashion in which emotions are experienced by listeners (Davies, 2001). Although the debate is at a theoretical level, its resolution has practical implications for interpreting music effects. Indeed, if music is only representative of emotion, its therapeutic value could be seriously questioned. Studies measuring physiological, endocrine and brain responses to music as indices of emotional reactivity have supported the emotivist view, but the nature of these emotional responses and their resemblance with emotions induced by other stimuli is unclear.

1.1. Autonomic nervous system responses

In order to show that people not only recognize but feel emotions in response to music, emotional reactions should be measured by techniques that are independent of voluntary subject control, such as psychophysiological measures. Following this line of research, Krumhansl (1997) compared the autonomic responses elicited by different musical emotions and found that sad, happy, and fearful music could be differentiated by their autonomic activation patterns: Sad music was most strongly associated with changes in heart rate, blood pressure, skin conductance and skin temperature, fearful music was mostly associated with changes in the rate and amplitude of blood flow, and happy music principally produced changes in respiratory activity and showed the highest skin conductance level (SCL). However, subsequent studies have failed to replicate many of these findings. Khalfa et al. (2002) found that skin conductance responses (SCR) were highest during the listening of fearful music, Baumgartner et al. (2006) observed increased SCL during sad and fearful music compared to happy music, and Nater et al. (2006) found higher SCL during the listening of unpleasant compared to pleasant music. Moreover, Nater et al. (2006) found higher heart rates during unpleasant compared to pleasant music, whereas Sammler et al. (2007), Witvliet and Vrana (2007), and Krumbansl (1997) found the opposite. Therefore, there are inconsistent findings of the intensity and direction of these autonomic responses between studies.

Such inconsistencies across psychophysiological emotion studies are relatively common (Cacioppo et al., 2000), and the outcomes may be related to some context-bound patterns of actions that allow the same emotion to be associated with a wide range of behavior and varying patterns of somato visceral activation (Lang et al., 1990). However, it should be noted that some psychophysiological measures appear more reliable than others. For example, respiration rate appears to be consistently higher during happy and fearful music than during...
Brain imaging techniques provide yet another way to measure emotional reactions objectively. Studies using such techniques have shown that pleasant emotional reactions to music activate regions previously known to be involved in approach-related behaviors, such as the prefrontal cortex (Blood and Zatorre, 2001; Blood et al., 1999; Koelsch et al., 2006; Menon and Levitin, 2005), periaqueductal gray matter (Blood and Zatorre, 2001), and the nucleus accumbens (Blood and Zatorre, 2001; Menon and Levitin, 2005). Negative emotions in contrast activate regions involved in withdrawal-related behavior, such as the parahippocampal gyrus (Blood et al., 1999) and amygdala (Koelsch et al., 2006). Although these observations are fairly consistent with activations observed with other emotional inducers, brain activations alone do not allow for the distinction between processes involved in emotional perception and emotional feeling. Physiological changes that affect the body and its responses are necessary to demonstrate the induction of emotional feelings.

1.4. Present study

Although these studies demonstrate that some emotions are felt in response to music, the results do not definitively refute the cognitivist viewpoint, as many psychophysiological responses are inconsistent, and the responses that appear to induce the most stable responses (e.g., respiration rate or hormonal responses) may be influenced by other confounding factors, such as arousal or distraction. Finally, brain imaging techniques cannot solely discriminate emotional feelings from other aspects of emotional processing.

In order to demonstrate the induction of emotional feelings, involuntary changes that affect the body and emotional processing have to be observed in response to musical excerpts conveying different emotions. In this context, the startle reflex is a good candidate measure, as it has been extensively and successfully used to probe emotional reactions. It is an automatic defensive reaction to surprising stimuli and can be measured by the magnitude of the eye blink triggered by a loud white noise. As a response of the defensive emotional system, it is frequently used to test the efficacy of anxiolytic drugs (Winslow et al., 2007) or to explore emotional reactivity in affective disorders (Grillon and Baas, 2003). In normal individuals, it is typically enhanced by negative emotions and diminished by positive ones, using pictures (Lang et al., 1998), films (Kaviani et al., 2004), or sounds (Bradley and Lang, 2000) to induce emotions. The present study applied an affective startle modulation paradigm to musical stimuli and compared the effects of pleasant and unpleasant musical excerpts on the acoustic startle blink reflex. If emotions are induced during music listening, then the startle reflex should be larger and of shorter latency during unpleasant music compared to pleasant music.

Moreover, in order to measure music effects on emotional reactions, heart rate and skin conductance responses were also obtained along with facial expressions by assessing electromyographic (EMG) activity of the zygomaticus major (smiling) and the corrugator supercilii (frowning). Previous studies have shown that the activity of these muscles discriminated well between pleasant and unpleasant emotions elicited by pictures (Lang et al., 1998). Thus, it was expected that zygomatic activity would be higher during pleasant music, and corrugator activity to be more noticeable during unpleasant music (Witvliet and Vrana, 2007).

2. Methods

2.1. Participants

Sixteen participants (9F, 7M), aged between 20 and 40 years (M=25.1±9.3 years) took part in this study. None were musicians, all reported fewer than five years of musical training, and none claimed any regular practice of a musical instrument.

2.2. Musical excerpts

The musical excerpts used in this study were adapted from a prior study on pain modulation (Roy et al., 2008). Three 100 s excerpts of pleasant music and three 100 s excerpts of unpleasant music were selected from a pool of 30 musical excerpts. Each of the 30 excerpts had been previously evaluated by 20 independent participants on the dimensions of valence (on a 0=‘pleasant’ and 9=‘unpleasant’) and arousal (with 0=‘relaxing’ and 9=‘stimulating’). Three highly pleasant and three highly unpleasant excerpts were selected. Since unpleasant excerpts were always judged to be arousing, all excerpts were selected in the high range of arousal. Pleasant excerpts were judged to be more
pleasant than unpleasant excerpts (mean valence for pleasant excerpts=2.40, mean valence for unpleasant excerpts=6.68; t(19)=5.58, p < 0.05) and did not differ in arousal (mean arousal for pleasant excerpts=5.00, mean arousal for unpleasant excerpts=5.18; t (19)=1.535, n.s.). The selected pleasant excerpts were taken from the classical or jazz/pop repertoire and could be described as uplifting, with a rather fast tempo, such as the “Opening of William Tell” by Rossini. Unpleasant excerpts were mainly taken from the contemporary music repertoire. Examples of excerpts for each emotion category can be heard on our web site at www.brams.umontreal.ca/peretz. All selections were normalized to equate loudness across musical excerpts by setting the peaks of the excerpts at 8% of the maximum volume allowed, using the normalisation option of the Cool Edit 2 sound editing software.

The primary emotions (sadness, happiness, fear, anger, peacefulness and surprise) and moods (anger, depression, fatigue, anxiety, vigor and confusion, as measured with the “profile of mood states”, POMS; McNair et al., 1992), induced by those excerpts were also previously assessed (Roy et al., 2008). Results showed that the primary emotions conveyed by the excerpts were consistent with their emotional valence. Pleasant excerpts were associated with happiness whereas unpleasant ones were associated with fear and anger. Results on the mood questionnaire confirmed these observations: The subscales for highly arousing negative moods, such as anger and anxiety, were higher after listening to the unpleasant excerpts, whereas the subscales for less arousing moods, such as depression, remained unaffected. Thus, the selected excerpts convey primary emotions and induce moods consistent with their positive or negative valence and their high level of arousal.

2.3. Data collection and reduction

Startle responses were elicited by a 100 dB SPL, 50 ms burst of white noise, with instantaneous rise time. The acoustic startle probe was presented over Sony MDR-v200 headphones. The eye blink component of the startle reflex was recorded electromyographically from the orbicularis oculi muscle beneath the left eye, using two miniature 4 mm Ag/AgCl shielded electrodes placed 1.5 cm apart and a signal ground electrode placed on the forehead, following the guidelines of Blumenthal et al. (2005). The signal was amplified by 1000 and band-pass filtered at 90 Hz–500 Hz using a Biopac MP150 System (Biopac Systems, Inc., Santa Barbara, CA). The sampling rate was set at 1000 Hz. The amplified signal was then transformed using the root mean square.

The maximum amplitude and latency of each startle response were extracted from the data. Following the guidelines of Balaban et al. (1986), only responses for which the onset occurred between 21 and 120 ms from noise onset were considered as startle responses and included in the analysis. The raw blink measurements were then standardised within each subject to decrease variability due to differences in the absolute size of the startle blink across subjects, and expressed as T scores (50+10z), which yielded a mean of 50 and a standard deviation of 10 for each subject. The blink amplitudes and latencies T scores were then averaged for the pleasant and unpleasant music condition.

To assess the sound intensity of the musical excerpts prior to each startle probe, the total root mean square amplitude (RMS) power was extracted for 1 s windows preceding each burst of white noise and averaged for each subject and musical condition. Facial EMG was recorded over the left corrugator and zygomatic sites (Fridlund and Cacioppo, 1986), using 8 mm Ag/AgCl shielded electrodes. Signals were bandpass filtered from 90 Hz to 1000 Hz and transformed using the root mean square. Sampling rate was set at 1000 Hz. Area under the curve of the rectified EMG signal were then extracted for the corrugator and zygomatic muscle.

Electrocardiogram (ECG) was recorded using a standard 3 lead montage (Einthoven lead 2 configuration) (Biopac EL503). Instantaneous intervals between each R-wave of the ECG (RR) were calculated from the ECG using a peak detection algorithm to detect successive R-waves and obtain a continuous R–R tachogram. Careful examination of the ECG and the tachogram ensured that the automatic R-wave detection procedure had been performed correctly.

Skin conductance level (SCL) was recorded on the palmar surface of the left hand, at the thenar and hypothenar eminences (Fowles et al., 1981). The signal was smoothed and the mean SCL was calculated for the whole duration of each musical excerpt and averaged for the pleasant and unpleasant music condition.

2.4. Procedure

The physiological sensors were affixed while the participants sat comfortably in a quiet room. The pleasant and unpleasant excerpts were presented in a counterbalanced order across participants. Fig. 1 illustrates the procedure for one musical excerpt. Each musical excerpt started with an emotional induction period of 21.3 s in which there were no startle probes. The remaining 78.7 s were divided in six 11 s time window in which a startle probe occurred randomly. Each time window was separated by a period of 2.3 s in which no startle probe occurred. After each musical excerpt, the subject rated his/her emotional reaction to the music on the dimensions of valence (0=unpleasant, 9=pleasant) and arousal (0=relaxing, 9=stimulating).

2.5. Data analysis

The musical excerpts were assessed statistically for the expected emotional effects by comparing the mean ratings of valence and arousal. Sound intensity was assessed before each startle probe in both musical conditions using RMS power preceding each startle probe for the pleasant and unpleasant music conditions. After these control analyses, a multivariate analysis of variance (MANOVA) was conducted to test if some patterns of physiological activation could reliably discriminate between the pleasant and unpleasant musical conditions.

3. Results

3.1. Self-reported emotions

The mean valence and arousal ratings were calculated for the pleasant and unpleasant excerpts. The t-tests performed on these
average ratings confirmed that the intended emotions of the musical excerpts were well recognized. The pleasant and unpleasant excerpts differed significantly on the dimension of valence (with a mean rating of 8.49 and 1.91, respectively; \( t(15)=13.04, p<0.001 \)). In contrast, pleasant and unpleasant musical excerpts did not differ on the dimension of arousal (with a mean rating of 6.41 and 7.50, respectively; \( t(15)=1.79, \text{n.s.} \)).

### 3.2. Sound amplitude

The RMS power of the 1 s windows preceding each startle probe was equivalent for the pleasant (23.03±0.68) and unpleasant (23.21±0.49) musical excerpts (\( t(15)=0.31, \text{n.s.} \)).

### 3.3. Physiological measures

Fig. 2 illustrates the mean values of each physiological measure for the pleasant and unpleasant music condition. The results of the MANOVA showed that the physiological responses were significantly affected by the musical condition (\( F(6, 10)=6.86, p<0.01, \eta^2=0.81 \)). The startle blink reflex was larger (\( t(15)=3.35, p<0.01, \eta^2=0.43 \)) and faster (\( t(15)=2.81, p<0.05, \eta^2=0.35 \)) during the unpleasant music as compared to the pleasant music. Activity of the corrugator muscle was higher during the unpleasant condition (\( t(15)=2.79, p<0.05, \eta^2=0.34 \)), but no significant difference in the activity of the zygomatic muscle was obtained between the two musical conditions (\( t(15)=1.35, \text{n.s.}, \eta^2=0.11 \)). RRI was also not affected by the valence of the musical excerpts (\( t(15)=0.72, \text{n.s.}, \eta^2=0.03 \)). In contrast, the SCL was found to be larger during the pleasant than the unpleasant music condition (\( t(15)=2.43, p<0.05, \eta^2=0.28 \)). The physiological responses were not limited to the 21.3 initial seconds without startle probes but were considered for the whole excerpt instead because only non-significant trends in the same direction were obtained on the initial part of the musical excerpts.

### 3.4. Discriminant analysis

The results of the discriminant analysis showed that the pleasant and unpleasant excerpts could easily be differentiated by a single function (Wilk’s lambda (6)=0.51, \( p<0.01 \)). This function correctly classified pleasant excerpts in 75% of cases and unpleasant excerpts in 87.5% of cases. Table 1 summarizes the canonical variate correlation coefficients of each physiological variable for the discriminant function. These canonical variate correlation coefficients were much more important for startle amplitude and latency compared to the other physiological measures, indicating that the startle reflex was the measure that contributed the most to the separation of the musical conditions. Although there is a lack of consensus regarding how high correlations in a loading matrix should be interpreted, typically only

![Image](https://example.com/image.png)

**Fig. 2.** Mean physiological responses for the pleasant and unpleasant musical conditions. Error bar represents one standard error above the mean. Significant differences (\( p<0.05 \)) are indicated by asterisks. Note that error bars reflect the between-subjects variance in each condition whereas the results of the statistical tests reflect the within-subject contrast across experimental conditions.

### Table 1

<table>
<thead>
<tr>
<th>Measure</th>
<th>Canonical Variate Correlation Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Startle amplitude</td>
<td>0.875</td>
</tr>
<tr>
<td>Startle latency</td>
<td>-0.728</td>
</tr>
<tr>
<td>Corrugator activity</td>
<td>0.080</td>
</tr>
<tr>
<td>Zygomatic activity</td>
<td>-0.074</td>
</tr>
<tr>
<td>Skin conductance level</td>
<td>-0.030</td>
</tr>
<tr>
<td>R-R interval</td>
<td>-0.009</td>
</tr>
</tbody>
</table>

*Note: p-values and effect sizes are calculated based on the within-subject contrast across experimental conditions.*
variables with loadings of .32 and above are considered interpretable. 
Comrey and Lee (1992) suggest that loadings over 0.71 are considered 
excellent, 0.63, very good, 0.55, good, 0.45, fair and 0.32, poor. In 
light of those guidelines, the canonical variates coefficients appear 
excellent for the startle reflex, but difficult to interpret for the other 
physiological measures.

4. Discussion

4.1. Induction of emotional feelings by music

The startle reflex was of higher amplitude and shorter latency 
during the listening of unpleasant in comparison with pleasant 
excerpts, suggesting that different emotional states were effectively 
induced by music. As the musical excerpts were manipulated to vary 
on the dimension of valence independently of arousal or loudness, the 
observed effects are likely to reflect the induction of positive and 
negative emotional states in response to music, thereby supporting the 
emotivist’s stance in contrast to critiques of cognitivists. First, the 
startle reflex is an involuntary response that does not depend on the 
doubtful capacity of the subjects to adequately describe their own 
experience. Second, the affective startle modulation effect is a reliable 
measure that avoids the important variability characterizing auto-
nomic nervous system measures. Moreover, startle modulation can be 
ascribed to the emotional valence rather than arousal or attention, 
thereby contrasting with prior studies. Third, because the modulation 
of the startle reflex indicates facilitation or inhibition of a motivational 
propensity to withdraw, it convincingly distinguishes the induction of 
emotional states from the cold perception of emotional features. Thus, 
music appears to be as powerful as pictures (Lang et al., 1998), films 
(Kaviani et al., 2004) or natural sounds (Bradley and Lang, 2000) to 
induce positive and negative emotions.

The absence of a neutral control condition complicates the 
comparisons with studies using different inducers of emotion, since 
it is difficult to tell if the startle modulation was due to an increase of 
the reflex during unpleasant music, a decrease of the reflex during 
pleasant music, or a combination of both. The absence of a neutral 
control condition, however, was not incidental as it is difficult to find 
“neutral” music as judged by a majority of listeners. In addition, the 
use of a silent control condition would not have been informative as 
sound level by itself has been shown to influence the acoustic startle 
(Franklin et al., 2007). Similarly, white noise matched with the musical 
excerpts for sound amplitude could not have been a better control 
condition because white noise is generally experienced as unpleasant. 
Nonetheless, a study using a similar design in which three startle 
probes were delivered within 2-min long video clips, showed that 
pleasant videos reduced the magnitude of the startle reflex compared 
to neutral videos, whereas unpleasant ones increased it (Kaviani et al., 
2004). Hence, it is reasonable to assume that the present results 
depend on the effect of a combination of both facilitation and 
inhibition of the defensive system. Nevertheless, the fact that the 
startle reflex was modulated by the emotional valence of the excerpts 
is sufficient to attest that emotional states were induced by music.

4.2. Startle reflex and physiological recordings as indices of felt emotion

In support to the idea that the startle reflex is one of the most 
reliable indices of emotional valence (Lang et al., 1998), the startle 
reflex was the best response to discriminate between pleasant and 
unpleasant musical excerpts. Corrugator activity was the second most 
discriminative measure but was far behind the startle reflex. Although 
its contribution to the discriminant function was minimal, the analysis 
of variance showed that corrugator activity was significantly higher 
during the unpleasant excerpts compared to the pleasant excerpts, 
confirming that positive and negative emotions were felt during the 
listening of those excerpts. Zygomatic activity, however, was not 
shown to be significantly modulated by the valence of the excerpt. This 
lack of sensitivity for zygomatic compared to corrugator activity 
is a common finding (Larsen et al., 2003), perhaps because the 
zygomatic is implicated in some negative emotions such as disgust or 
that it is involved with display rules and other fine voluntary motor 
behaviors.

Skin conductance, while having little contribution to the discrimi-
nant function, proved to be higher during the pleasant compared to the 
unpleasant excerpts. This finding adds to the controversy surrounding 
the interpretation of skin conductance changes in response to musical 
emotion. The present findings are consistent with those of Krumhansl 
(1997) but inconsistent with Baumgartner et al. (2006) and Nater et al. 
(2006). Moreover, the present findings are opposite to those generally 
observed with other emotional induction techniques such as mental 
imagery or the presentation of emotional movie clips for which SCL is 
higher during negatively valenced emotion (Cacioppo et al., 2000). This 
outcome suggests that skin conductance levels might be related to 
some aspects of the emotional response that are not directly linked to 
the perceived valence and arousal and may vary from one study to 
another. In the present case, the pleasant and arousing excerpts might 
have prompted motoric activity such as dancing or tapping of the foot. 
Finally, no differences in heart rate were found between pleasant and 
unpleasant excerpts. Taken together with the negative findings of 
Baumgartner et al. (2006) and Etzel et al. (2006), this lack of difference 
suggests that heart rate alone is not sufficient to differentiate pleasant 
from unpleasant musical conditions.

4.3. Implications

The present study supports the emotivist stance and provides some 
thoretical justifications for the use of music as a therapy. If music is able 
to induce emotions that can reduce the activity of the emotional 
defensive system, it can be used to alleviate some unpleasant 
emotional states, such as anxiety (Rudin et al., 2007), depression 
(Siedlecki and Good, 2006), or pain (Roy et al., 2008). In addition, the 
demonstration that emotions are indeed felt in response to music 
also opens up questions about how and why it does so. The combination 
of psychophysiological recordings with brain imaging 
techniques, in addition to self-reported measures of emotion 
and careful manipulation of the musical stimuli, will help to characterize 
how the brain and the body interact to create emotional feelings to 
music (Koelsch, 2005).

Acknowledgements

The work was supported by a grant from the Natural Science and 
Engineering Research Council of Canada (NSERC) to Isabelle Peretz, 
and by a doctoral scholarship from the NSERC to Mathieu Roy. We 
thank Amee Baird for English editing, Francine Giroux for statistical 
advice and Pierre Rainville for his suggestions on physiological signal 
analysis.

References

Balaban, M.T., Losito, B.D., Simons, R.F., Graham, F.K., 1986. Off-line latency and 
amplitude scoring of the human reflex eye blink with Fortran IV. Psychophysiology 
23, 612–621.
Baumgartner, T., Esslen, M., Jancke, L., 2006. From emotion perception to emotion 
60, 34–43.
98, 11818–11823.
and unpleasant music correlate with activity in paralimbic brain regions. Nat.
Neurosci. 2, 382–387.
Committee report: guidelines for human startle eyelink electromyographic 