Available online at www.sciencedirect.com

SCIENCE DIRECT®

International Journal of Psychophysiology 53 (2004) 91–103

INTERNATIONAL JOURNAL OF PSYCHOPHYSIOLOGY

www.elsevier.com/locate/ijpsycho

Affective and physiological responses to environmental noises and music

Patrick Gomez, Brigitta Danuser*

Institute for Hygiene and Applied Physiology, Swiss federal Institute of Technology, Clausiusstrasse 25, 8092 Zurich, Switzerland

Received 3 September 2003; received in revised form 12 February 2004; accepted 18 February 2004
Available online 21 April 2004

Abstract

Research suggests that respiratory patterns may reflect general dimensions of emotional response. In this study, we investigated the relationships between judgments of affective valence (pleasantness) and arousal and respiratory responses to acoustic stimuli. Sixteen environmental noises and 16 musical fragments of 30 s duration were presented to 31 participants, while respiration, skin conductance level and heart rate were recorded. Judgments of valence and arousal were registered using the 9-point Self-Assessment Manikin. For noises, breathing accelerated and minute ventilation augmented with decreases in pleasantness for low-arousal stimuli and with increases in arousal for positive stimuli. For music, breathing accelerated and minute ventilation augmented with increases both in rated valence and arousal. Skin conductance level increased with arousal ratings for music but not for noises, whereas mean heart rate increased with rated arousal for noises but not for music. Although both noises and music are sound-vibrations, differences in the relationships between affective judgments and physiological responses were found suggesting differences in the processing of the two types of acoustic stimuli.

Keywords: Arousal; Heart rate; Music; Noises; Respiration; Skin conductance level; Sounds; Valence

1. Introduction

Emotional reactions can be evoked by stimuli in different sensory modalities. Those associated with processing pictures have been extensively investigated and several studies have provided evidence for a correspondance between physiological variables (skin conductance, heart rate, startle blink reflex, facial EMG, brain activity) and the emotional dimensions of valence and arousal (see Hamm et al., 2003 for review). We recently investigated respiratory responses while viewing affective pictures using the same dimensional framework (Gomez et al., 2004a). Several respiratory measures were found to covary with judgments of valence and arousal. To extend our knowledge of respiratory responses during affective processing, we investigated in this study affective and respiratory reactions while listening to environmental noises and musical passages.

Both noises and music are sound-vibrations. However, one would be hard-pressed to call the crying of a baby, a siren or a ringing telephone 'music'. Thus,

* Corresponding author. *Present address*: Patrick Gomez and Brigitta Danuser: Institut universitaire romand de Santé au Travail, Rue du Bugnon 19, 1005 Lausanne, Switzerland. Tel.: +41-21-314-74-21; fax: +41-21-314-74-20.

E-mail address: brigitta.danuser@hospvd.ch (B. Danuser).

0167-8760/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijpsycho.2004.02.002

differences must exist between what we normally call noises and what we normally call music. These differences may be associated with different processing of the two types of acoustic stimuli and reflect themselves in different physiological responses. Theories of auditory event perception have highlighted a distinction between 'everyday' and 'musical' listening (Gaver, 1993). However, this account of listening in two ways has been recently challenged by Dibben (2001).

Only one study has assessed physiological responding (skin conductance, heart rate, startle blink reflex, facial EMG) to a large sample of naturally occurring sounds (Bradley and Lang, 2000). To our knowledge, no study has yet explored the relationships between self-reported affective responses and respiratory reactions to natural noises. A number of studies have investigated affective and physiological responses to music (for review see Bartlett, 1996). In her doctoral thesis, Witvliet (1998) adopted the framework of valence and arousal to assess the effects of music on skin conductance, heart rate, facial EMG, and the startle reflex. Few studies, however, have assessed respiratory responses (Krumhansl, 1997; Nyklícek et al., 1997) and none have included a detailed breathing analysis as recommended by respiratory psychophysiologists (Wientjes, 1992; Boiten et al., 1994).

The available evidence concerning the relation between emotions and respiration indicates that more rapid breathing is associated with increases in arousal (e.g. Nyklícek et al., 1997; Boiten, 1998; Van Diest et al., 2001). Until now, no clear relationship between respiratory volume and emotional dimensions have been obtained (Boiten, 1998; Gomez et al., 2004a,b). Mean inspiratory flow and minute ventilation increase with increases in arousal (Boiten et al., 1994; Boiten, 1998; Gomez et al., 2004a) but may also be influenced by the affective valence of the stimuli (Gomez et al., 2004a). Finally, very little is known about the influence of affective stimuli on the relative contribution of the rib cage and the abdomen to respiratory volume (Boiten, 1998).

The main purpose of the present study was to assess the relationships between judgments of affective valence and arousal and physiological responses to noises and to music. To this end, we presented 16 noise sources and 16 musical fragments to 31 participants, each lasting 30 s in duration. We recorded respiration, skin conductance level, heart rate and judgments of valence and arousal. Moreover, we were interested in comparing the two-dimensional distribution of valence and arousal ratings for noises and musical fragments.

2. Methods

2.1. Participants

Participants were 16 men and 15 women, mostly university students, ranging in age from 18 to 37 years (mean age 24 years). They reported no long-term hearing impairment and were healthy on the day of testing. They were asked to avoid drinking coffee and alcoholic beverages prior to the experiment.

2.2. Affective stimuli

Participants listened to 16 noises and 16 musical fragments with a duration of 30 s each. The 32 stimuli were selected from a larger sample of 25 noises and 25 music excerpts based on evaluation in pre-testing in our laboratory. The selected stimuli were expected to engage a broad range of emotional responses, varying widely in affective valence and arousal. We paid particular attention to select both noises and musical fragments with little variations in their characteristics over the 30-s presentation, in order to guarantee high intra-stimulus homogeneity with respect to their levels of valence and arousal. All selected noises represented environmental acoustic stimuli. The musical excerpts were all instrumental and were drawn from the corpus of Western music. The 16 noises and the 16 musical selections are listed in the legends of Tables 1 and 2, respectively. The stimuli were edited digitally on a CD-ROM. The last second of each excerpt was faded out so as not to give the stimulus a clipped sound at the end. Sound intensities (Leq) at presentation ranged for noises from a low of 52.2 dB(A) (people playing tennis) to a high of 76.7 dB(A) (siren) and for music from a low of 58.9 dB(A) ('Adagio assai', M. Ravel) to a high of 77.5 dB(A) ('Black Arrows', Manowar). We did not match the stimuli for sound intensity, because the primary goal was to use ecologically valid sounds that effectively communicate affect (see Bradley and

Table 1
Valence and arousal ratings of the environmental noises. Means and S.D.s (in parentheses) and distribution in the two-dimensional affective space

Noise	Valence	Arousal	Pos. V high A	Pos. V mod. A	Pos. V low A	Neutral V high A	Neutral V mod. A	Neutral V low A	Neg. V high A	Neg. V mod. A	Neg. V low A
1	6.29 (2.30)	7.64 (1.25)	**15**	2	0	8	2	0	2	2	0
2	5.58 (2.17)	6.06 (2.00)	4	4	5	5	4	1	**6**	2	0
3	5.16 (1.75)	3.71 (1.97)	2	0	4	1	9	**10**	0	1	4
4	7.81 (1.56)	2.64 (1.50)	0	5	**20**	0	3	2	0	0	1
5	5.55 (2.05)	3.93 (1.75)	0	5	5	3	**7**	5	0	3	3
6	2.77 (2.01)	7.87 (1.43)	1	1	0	5	3	0	**20**	1	0
7	3.90 (1.97)	6.48 (1.48)	5	1	0	1	7	0	**11**	5	1
8	2.74 (1.55)	5.84 (2.05)	0	1	0	0	4	3	**13**	8	2
9	4.26 (2.25)	4.81 (2.69)	1	0	4	5	**6**	4	5	2	4
10	5.19 (1.85)	4.03 (2.17)	2	1	6	0	8	**10**	3	1	0
11	1.97 (1.30)	6.71 (2.51)	0	0	0	0	1	2	**21**	3	4
12	2.48 (1.75)	6.55 (1.75)	1	0	1	0	2	0	**16**	11	0
13	3.13 (1.46)	5.87 (1.73)	0	0	0	2	6	4	**11**	8	0
14	4.00 (2.57)	7.03 (1.92)	5	1	0	4	4	2	**12**	3	0
15	4.36 (2.20)	5.84 (1.90)	4	2	2	4	3	1	**8**	5	2
16	4.71 (2.30)	6.19 (2.09)	5	1	2	6	4	0	**7**	4	2

Notes: The selected noises were: 1 cheering spectators at sport event, 2 sea roar, 3 people playing tennis, 4 little stream with bird twitter, 5 chirping crickets, 6 siren, 7 crowded city center, 8 ringing telephone, 9 water-drops, 10 ringing bell, 11 pneumatic hammer, 12 highway, 13 dentist, 14 flying airplane, 15 waterfall, 16 storm. Pos.: positive; neg.: negative; mod.: moderate; V: valence; A: arousal; neg./low: values 1–3; neutral/mod.: values 4–6; pos./high: values 7–9; bold is the area receiving the largest number of ratings.

Table 2
Valence and arousal ratings of the musical selections. Means and S.D.s (in parentheses) and distribution in the two-dimensional affective space

Musical selection	Valence	Arousal	Pos. V high A	Pos. V mod. A	Pos. V low A	Neutral V high A	Neutral V mod. A	Neutral V low A	Neg. V high A	Neg. V mod. A	Neg. V low A
1	6.55 (2.41)	7.61 (1.69)	**17**	2	0	6	0	1	4	0	1
2	6.84 (2.27)	7.61 (1.43)	**21**	0	0	4	3	0	2	0	1
3	7.07 (1.97)	6.48 (2.20)	**18**	3	2	0	3	2	1	2	0
4	7.13 (1.34)	6.00 (1.59)	**11**	9	1	2	7	1	0	0	0
5	6.13 (2.50)	3.23 (2.03)	2	4	**12**	0	1	5	1	4	2
6	5.94 (1.86)	2.77 (1.38)	1	2	**12**	0	2	10	0	1	3
7	5.29 (1.66)	6.71 (1.87)	6	1	2	**15**	2	2	1	2	0
8	5.23 (1.80)	5.61 (2.04)	3	3	2	**8**	7	4	2	1	1
9	6.35 (2.20)	2.90 (1.68)	1	4	**16**	0	1	3	0	1	5
10	5.81 (2.23)	2.87 (1.41)	1	3	**11**	0	3	7	0	0	6
11	6.61 (1.99)	2.81 (1.78)	1	4	**16**	1	1	5	0	0	3
12	3.87 (3.14)	8.29 (1.10)	9	1	0	3	0	0	**17**	1	0
13	3.19 (1.74)	6.48 (1.67)	2	0	0	3	4	0	**14**	6	2
14	4.35 (2.23)	4.19 (2.07)	0	3	3	3	3	**5**	4	**5**	**5**
15	5.06 (2.84)	2.74 (1.86)	2	1	**11**	0	2	3	0	1	**11**
16	5.06 (2.62)	1.97 (1.38)	0	1	10	0	0	7	1	1	**11**

Notes: The musical selections were: 1 Offenbach – Rosenthal, Cancan, 2 A. Ponchielli, La Gioconda – Dance of the Hours, 3 E. Serra, Le Grand Bleu – Spaghetti del mare, 4 G. Mahler, Symphony No. 1 – langsam schleppend, 5 E. Elgar, Salut d' Amour, 6 G. Holst, The Planets – Venus, 7 R. Strauss, Also sprach Zarathustra – Von den Freuden und Leidenschaften, 8 E. Serra, Le Grand Bleu – The Monastery of Amorgos, 9 A. Dvorák, Symphony No. 9-Largo, 10 E. Elgar, Enigma Variations – Romanza, 11 Gandalf, From Source to Sea – Refuge Island, 12 Manowar, Hail to England – Black Arrows, 13 S. Barber, Adagio for strings Op. 11, 14 D. Borgir, Puritanical Euphoric Misanthropia – Fear and Wonder, 15 G. Mahler, Symphony No. 5 – Adagietto, 16 M. Ravel, Piano Concerto in G major – Adagio assai. Pos.: positive; neg.: negative; mod.: moderate; V: valence; A: arousal; neg./low: values 1–3; neutral/mod.: values 4–6; pos./high: values 7–9; bold is the area receiving the largest number of ratings.

Lang, 2000). Presentation was controlled using an IBM computer connected with a JVC Ax-Z711 amplifier. Sounds were presented over a pair of Revox Plenum B MK II speakers.

The presentation order was determined as follows. The 16 noises and the 16 musical excerpts were ranked separately from low to high using the formula 'mean valence+mean arousal' of the ratings obtained in the pre-testing and were divided into four groups (ranks 1–4, 5–8, 9–12 and 13–16). Then, four blocks of eight stimuli each, including four noises and four musical excerpts were formed by drawing by lot a stimulus from each group. Of the 24 possible arrangements of the four blocks, eight were randomly selected, with the condition that each block had to be twice in first, second, third and fourth position. The order of the eight stimuli within each block was randomly determined and was different for the eight arrangements. The eight presentation orders were counterbalanced across subjects. Thus, the noise and musical stimuli were not presented separately but were mixed together. This was done to guarantee a better comparability between noises and musical excerpts as regards affective ratings.

2.3. Affective and physiological response measurements

Judgments of valence and arousal were registered using the pencil-and-paper version of the 9-point Self-Assessment Manikin (SAM; Bradley and Lang, 1994).

Respiration was measured using a respiratory inductive plethysmograph (Respitrace PLUS, SensorMedics, NIMS, USA). The coils of the Respitrace were placed on the chest 2 cm below the nipple line (for women this was right below the breast) and on the abdomen at the umbilical level, providing the thoracic motion (Vth) and the abdominal motion (Vab) signals. The proportionality constant between Vth and Vab amplifiers was determined by the qualitative diagnostic calibration method (QDC) carried out during a 5-min period of spontaneous breathing (Sackner et al., 1989). The Respitrace signals (in volts) were recorded with a Labview program (National Instruments, USA) and analyzed by a MATLAB program (MathWorks Inc., USA). Volume calibration of both channels was performed before and after the experiment (pre- and post-calibration) with a calibrated pneumotachograph (CPX/D, Medical Graphics Corporation, USA), employing a procedure in which the subject breathed through a mouthpiece at different rates and depths (natural – relatively fast and shallow – relatively slow and deep – natural) for approximately 20 s for each phase. A multiple linear regression was performed over the calibration periods, yielding the regression coefficient to be used for the reconstruction of the volume out of the Respitrace signals. Calibration was accepted when the overall regression factor r^2 was over 0.8. The calculated regression coefficient was implemented into the MATLAB program, which analyzed the breathing curves and arranged the data in a log-file. Both the Respitrace and the pneumotachograph signals were sampled at 200 Hz.

For each breath, the following respiratory parameters were extracted: inspiratory time (Ti) and expiratory time (Te, including expiratory pause), inspiratory volume (Vi) and expiratory volume (Ve), and the percentage of rib cage contribution to Vi (%RC). Total breath duration (Ttot), inspiratory duty cycle (Ti/Ttot), mean inspiratory flow (Vi/Ti) and minute ventilation (MV, inspiratory volume/total breath duration×60 s min^{-1}) were calculated off-line. Only Vi was used as a measure of tidal volume.

Skin conductance level (SCL) and heart rate (HR) were measured with the Varioport Measurement System (Becker Meditec, Karlsruhe, Germany), an 8-channel recording system. After amplification and filtering, data were digitized (12 Bit resolution) and saved on a compactflash-card. The device was controlled via computer running VP_Basic software (TEMEC Instruments B.V., 6460 HA Kerkrade, NL) and the channels were defined as follows: SCL: unit μS, sampling rate 256 Hz, saving rate 16 Hz, range 0.1–100 μS, resolution 0.001 μS; ECG: unit mV, sampling rate 512 Hz, saving rate 256 Hz, resolution 1 μV; Pulse: unit beats/min, sampling rate 512 Hz, saving rate 4 Hz, resolution 1 beat/min. Pulse was determined on-line from the ECG signal. SCL was recorded using a 16 Bit-resolution unit, which provided a constant 0.5 V across electrodes. Before recording, the palm of the non-dominant hand was cleansed with distilled water and two Ag/AgCl electrodes (4 mm contact area diameter, GE Medical Systems, Milwaukee, USA) filled with lubricating jelly (MED-TEK/Synapse, Arcadia, CA 91007, USA) were placed adjacently on the hypothenar eminence of the palm. For ECG recordings, pregelled Ag/AgCl electrodes (60 mm, Skintact®, Leonherd Lang GmbH, Innsbruck, Austria) were placed on the sternum and the left costal arch at the level of the 10th rib, with a reference below the right collar-bone.

2.4. Procedure

Participants were tested individually in one experimental session lasting approximately 1.5 h. The experiment took place in a sound-insulated, air-conditioned, dimly lit room. After arrival, participants filled out an informed consent form and the experimenter provided them with an outline of the experimental procedure. Participants were told that 16 noises and 16 musical passages of 30 s duration would be played in random order, and that each stimulus should be attended for the entire duration of presentation. They were then told that after each stimulus they had to report how they felt while listening to the noise or musical fragment. The importance of rating the stimulus as they actually felt while they listened to it was emphasized. Next, they were explained how to use the SAM rating scales and three practice stimuli were played in order to familiarize them with the SAM rating procedure. They were also instructed during the entire experiment to minimize all body movements so as not to disturb the physiological measurements. The participants sat in a comfortable armchair, in front of a table leaning on the back of the seat, with both feet on the floor and the arms lying on the table. At approximately 2 m from them was a white screen under which the two loudspeakers were placed.

Following attachment of bands and sensors, the QDC-calibration and the pre-calibration with the pneumotachograph were performed. Five seconds before each stimulus, participants were verbally informed by the experimenter that the next stimulus would be displayed. The experimenter sat outside the experimental room and could communicate with the participants through the ajar door. After offset of each stimulus, participants immediately rated it in the dimensions of valence and arousal and then relaxed till the next stimulus was announced. Duration of the inter-stimulus period was 65 s. After the last rating, post-calibration was carried out. Following the removal of bands and sensors, participants completed a questionnaire upon their musical preferences, listening habits and musical training. In the end, they were thanked, paid for their participation and asked not to divulge any of the details of the experiment to other potential participants.

2.5. Data reduction

Our main interest was the physiological response when the individual has adapted to the affective stimulus. Therefore, we concentrated our analyses on the second half of stimulus presentation. For the breathing parameters, median values of the 15-s interval before stimulus announcement (baseline) and median values of the 15-s interval before stimulus offset (stimulus interval) were computed for each subject and each stimulus. These intervals were calculated from the end of the last full respiratory cycle before stimulus announcement and before stimulus offset, respectively. The beginning of the intervals fell always within a respiratory cycle. This cycle was entirely included in the intervals so that intervals were slightly longer than 15 s. On average, baselines had a mean duration of 17.3 ± 0.7 s and included on average 4.4 ± 0.6 respiratory cycles, whereas stimulus intervals had a mean duration of 17.0 ± 0.7 s and included 4.6 ± 0.9 respiratory cycles. Although 17 s are a relatively short time, they are sufficient to detect variations in breathing responses as previous studies have shown with even shorter presentation epochs and comparably long analyzed intervals (Ritz et al., 2002; Gomez et al., 2004a). In accordance with standard statistical practice and consistent with previous studies (e.g. Fridlund, 1991; Jäncke, 1994), the median values of the baselines and of the stimulus intervals were log-transformed to minimize skewness and heteroscedasticity. Then, response scores were calculated by subtracting the transformed baseline score from the transformed score of the stimulus interval. Similarly, SCL and HR were computed by subtracting the log-transformed median value of the 15-s interval before stimulus announcement from the log-transformed median value of the last 15 s of stimulus presentation. Log-transformation was necessary to accommodate the data so that assumptions of normality and linearity, as required by the statistical analysis adopted, were met. Medians instead of means were used because the median is a much more robust estimate of the center of a sample of data.

2.6. Statistical analysis

All statistics were calculated using Systat Version 10 for Windows (SPSS Inc., 2000). To assess the relationships between affective judgments and physiological response measures, mixed effects regression models (also known as multilevel regression analysis) with marginal maximum likelihood estimation were used. The procedure involves a combination of the EM algorithm and Fisher scoring. For details, see Hedeker and Gibbons (1996). This statistical approach was used in order to take into account the multilevel structure of the data (16 observations for noises and music nested within each subject). The models tested included a random intercept for each subject and fixed effects for valence rating (V), arousal rating (A) and the interaction term (V×A). Both models with and without the V×A interaction were tested. In the following, the two models are referred to as the V–A-interaction model and the V–A-only model, respectively. The significance level was set at 0.05.

Multilevel regression analysis assumes normality and linearity. Residuals inspection of the log-transformed data indicated a good conformity to these assumptions. However, some outlying observations still remained after transformation. They were removed as according to Systat's warning.[1] The number of deleted outliers is reported in the Section 3.

Mixed regression was used similarly to assess the correlation between valence ratings and arousal ratings by testing the following model for noises and musical selections separately: arousal rating = valence rating + (valence rating × valence rating) (linear and quadratic relation). Moreover, we were interested in determining if the affective ratings were more consistent for noises or musical fragments and if valence ratings were more consistent than arousal ratings. To this end, we calculated for each subject an 'index of differentiation' for the valence and arousal ratings of the noises and of the music excerpts. A 2 (stimulus type: noise vs. music excerpt)×2 (affective dimension: valence vs. arousal) analysis of variance (ANOVA) with repeated measures was carried out on the four indices. The indices were computed by subtracting each subject's rating from the mean across all subjects and by summing up the absolute values of these differences. The indices were further log-transformed. Pairwise comparisons were conducted using paired t-tests with Bonferroni correction to control for multiple testing.

[1] Under normal regression assumptions, studentized residuals have a t distribution. Systat looks at the t distribution and prints a warning when the t-value exceeds the expected 99% level, which corresponds to a value of approximately ± 3.

3. Results

3.1. Affective judgments

3.1.1. Range and distribution

Tables 1 and 2 present means and S.D.s of the affective ratings for the 16 environmental noises and the 16 musical selections, respectively. Acoustic stimuli varied dramatically both in rated valence and arousal. For noises, valence ranged from 1.97 (pneumatic hammer) to 7.81 (little stream with bird twitter), whereas arousal ranged from 2.64 (little stream with bird twitter) to 7.87 (siren). For music, valence ranged from 3.19 (Adagio for strings, S. Barber) to 7.13 (Titan, G. Mahler), whereas arousal ranged from 1.97 (Adagio assai, M. Ravel) to 8.29 (Black Arrows, Manowar). Moreover, for each subject the lowest and the highest rating for each SAM dimension was determined and means across subjects were calculated (intra-subject range). For noises, valence ranged from 1.26 to 8.45 and arousal ranged from 1.68 to 8.68; for music, valence ranged from 1.58 to 8.74 and arousal ranged from 1.42 to 8.68.

Tables 1 and 2 also show the distribution of ratings in the two-dimensional affective space. They report the number of subjects rating each noise and each musical excerpt in the nine areas defined by the combination of negative (ratings 1–3), neutral (4–6) and positive (7–9) valence with low (1–3), moderate (4–6) and high (7–9) arousal. For noises, there was a relatively high number of ratings in the negative valence high-arousal area and relatively few in the negative valence low-arousal area. For music, both the positive high-arousal and positive low-arousal excerpts were the most represented whereas the negative moderately arousing selections were the least represented.

Table 3
Means and S.D.s (in parentheses) of the indices of differentiation

Noises		Music	
Valence	Arousal	Valence	Arousal
1.379_a (0.130)	1.361_a (0.122)	1.447_b (0.111)	1.290_c (0.137)

Notes: Means having the same subscript are not significantly different at $P<0.05$ (Bonferroni adjusted).

3.1.2. Listener agreement

The 2×2 ANOVA revealed a significant effect of the affective dimension ($F_{(1,30)}=25.95$, $P<0.0001$), whereas the effect for the stimulus type was not significant ($F_{(1,30)}=0.01$, $P>0.9$). The interaction term was also significant ($F_{(1,30)}=15.54$, $P<0.001$). Means of the four indices of differentiation are reported in Table 3. The post-hoc analysis indicated that the arousal ratings of the music excerpts were more consistent than all other ratings (lowest mean index) and the valence ratings of the music excerpts were the least consistent (highest mean index), with the valence and arousal ratings of the noises falling in-between. We can recognize the particularity of the affective ratings of the music excerpts by looking at Table 2. As to the arousal ratings, we can see that several music excerpts were rated from the great majority of subjects in the same area (e.g. Cancan, Offenbach–Rosenthal; Dance of the Hours, A. Ponchielli; Black Arrows, Manowar; Adagio assai, M. Ravel). As to the valence ratings, on the contrary, several music excerpts had no real predominant area. In particular, three fragments (Black Arrows, Manowar; Adagietto, G. Mahler; Adagio assai, M. Ravel) showed a bi-polarization, i.e. they were rated positively and negatively by a similarly large number of participants. This was not the case for any noise sources.

3.1.3. Correlation between affective dimensions

The mixed regression analysis yielded both for the noises and the musical selections a significant valence effect and a significant valence × valence effect. The estimated coefficients were for noises: Intercept, 7.893 (±0.365; $P<0.001$); valence, −0.741 (±0.170, $P<0.001$); valence×valence, +0.042 (±0.017, $P<0.05$); for music: Intercept, 6.183 (±0.559; $P<0.001$); valence, −0.641 (±0.238, $P<0.01$); valence × valence, +0.061 (±0.022, $P<0.01$). Plots of the predicted values for these two models are represented in Fig. 1. Negatively valent noises tended to be rated as more arousing than positively valent noises (predominant negative linear trend), whereas for music positive and negative selections tended to receive higher arousal rating than neutral ones (predominant quadratic trend).

3.2. Relationships between affective judgments and physiological responses

3.2.1. Noise stimuli

The estimated models for all physiological measures are presented in Table 4. For Ttot and MV the V–A-interaction model revealed a significant valence effect and a significant V×A interaction. For Ti and Te, the valence effect and the interaction term approached the significance level ($0.05<P<0.1$). For Ttot the estimate of the valence effect was positive and that of the interaction term was negative. The combined effect of the two estimates is represented in

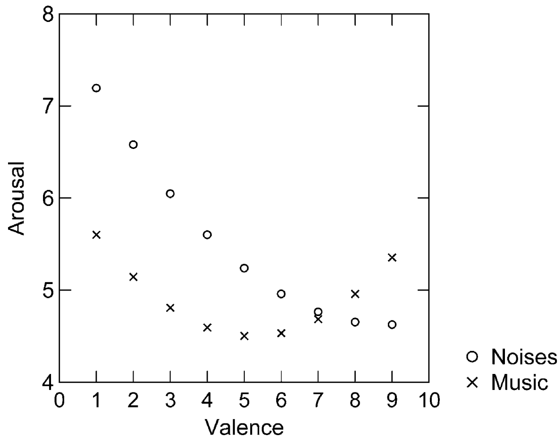

Fig. 1. Correlation between valence and arousal ratings. The estimated models represented are for noises: Arousal = 7.893 − 0.741 × valence + 0.042 × (valence × valence) and for music: Arousal = 6.183 − 0.641 × valence + 0.061 × (valence × valence).

Table 4
Estimated models for the physiological variables

Dependent variable	Noises			Music		
	predictor	estimate	se	predictor	estimate	se
Ti	V	8.17	4.60	**V**	**−6.02****	**1.89**
	A	2.53	4.07	**A**	**−4.84****	**1.70**
	V×A	−1.25	0.71			
Te	V	6.99	3.85	**V**	**−4.38****	**1.68**
	A	3.20	3.40	**A**	**−8.02******	**1.51**
	V×A	−0.80	0.59			
Ttot	**V**	**9.26****	**3.37**	**V**	**−5.18*****	**1.42**
	A	4.15	2.98	**A**	**−6.38******	**1.27**
	V×A	**−1.22***	**0.52**			
Ti/Ttot	V	−0.19	1.32	V	0.10	1.25
	A	−1.46	1.37	A	0.61	1.13
Vi	V	−0.51	1.27	V	0.36	1.33
	A	0.94	1.33	A	−0.80	1.20
Vi/Ti	V	−2.87	2.29	**V**	**4.71***	**2.07**
	A	**5.34***	**2.39**	**A**	**4.28***	**1.88**
MV	**V**	**−12.73****	**4.11**	**V**	**5.77*****	**1.65**
	A	−4.35	3.63	**A**	**4.81****	**1.50**
	V×A	**1.39***	**0.63**			
%RC	V	2.37	2.88	V	−1.32	0.82
	A	4.14	2.95	**A**	**1.76***	**0.74**
SCL	V	−0.15	0.63	V	0.73	0.62
	A	0.95	0.66	**A**	**3.03******	**0.57**
HR	V	0.32	0.46	V	0.20	0.44
	A	**1.18***	**0.48**	A	0.61	0.40

Notes: Ti: inspiratory time; Te: expiratory time; Ttot: total breath duration; Ti/Ttot: inspiratory duty cycle; Vi/Ti: mean inspiratory flow; MV: minute ventilation; %RC: percent rib cage; SCL: skin conductance level; HR: heart rate; V: valence rating; A: arousal rating; V×A: valence rating×arousal rating; se: standardized error. Bold: effects for V, A, and V×A with $P<0.05$; $*P<0.05$; $**P<0.01$; $***P<0.001$; $****P<0.00001$. The number of outlying observations removed were for noises: Ti: 1; Te: 2; Ttot: 1; Ti/Ttot: 4; Vi: 4; Vi/Ti: 0; MV: 0; %RC: 0; SCL: 0; HR: 3; for music: Ti: 0; Te: 1; Ttot: 5; Ti/Ttot: 0; Vi: 4; Vi/Ti: 1; MV: 0; %RC: 2; SCL: 4; HR: 0.

Fig. 2 (upper left) using a so-called two-dimensional contour plot. This is a plot with three variables; the *x* variable is valence and the *y* variable is arousal both ranging from 1 to 9; gradation lines represent different values of the *z* variable (physiological measure). The plot shows that Ttot tended to shorten with increasing arousal for positively valent noises, and to lengthen with increasing valence for low-arousal noises. For MV the estimate of the valence effect was negative and that of the interaction term was positive. As shown in Fig. 2 (upper right), MV increased with arousal for positive stimuli and decreased with valence for relatively low-arousal stimuli. Vi/Ti showed a significant positive correlation with arousal ratings. HR was higher while listening to high-arousal noises than to low-arousal noises (positive estimate). For Ti/Ttot, Vi, %RC, and SCL, nor the V–A-interaction model nor the V–A-only model indicated significant covariations with valence or arousal ratings. For these parameters, the V–A-only model is reported.

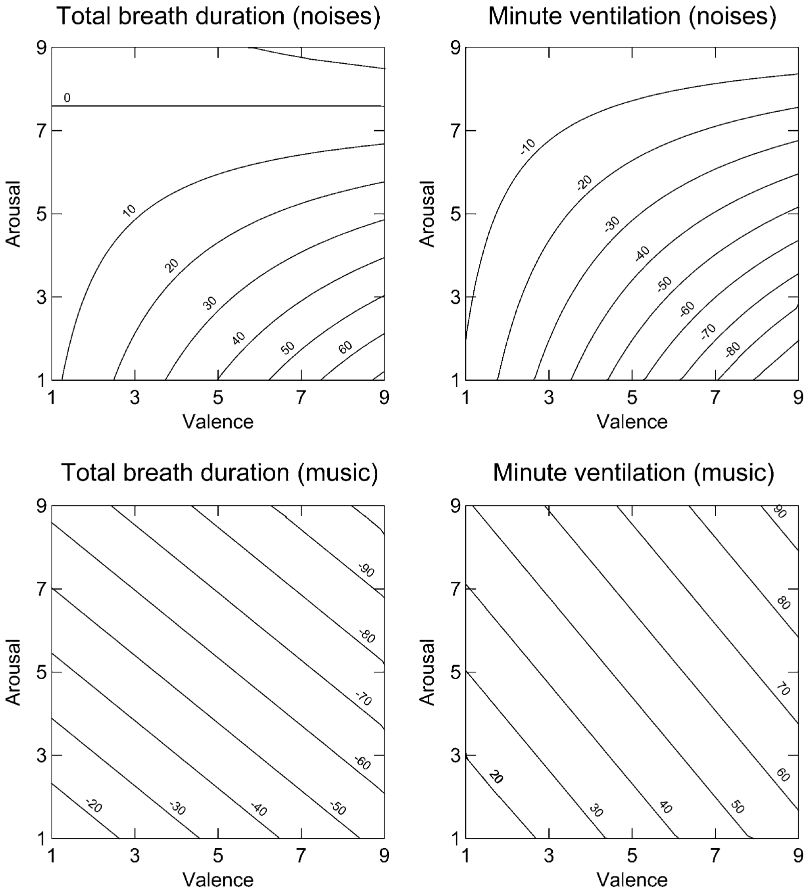

Fig. 2. Estimated models for total breath duration Ttot and minute ventilation MV for noises (upper row) and for music (lower row). *x*- and *y*-axes: 1 represents most negative valence/lowest arousal; 9 represents most positive valence/highest arousal. The estimated models represented are for noises: Ttot = 9.26 × valence − 1.22 × valence × arousal, MV = −12.73 × valence + 1.39 × valence × arousal; for music: Ttot = −5.18 × valence − 6.38 × arousal, MV = 5.77 × valence + 4.81 × arousal. Data are (log(stimulus interval) − log(baseline)) × 1000.

3.2.2. Musical stimuli

The estimated models are presented in Table 4. For Ti, Te, Ttot, Vi/Ti and MV both a significant valence and a significant arousal effect were obtained. Ti, Te and Ttot were shorter during positive than during negative music fragments (negative estimate) and during high-arousal stimuli than low-arousal stimuli (negative estimate) (Fig. 2, lower left). Vi/Ti and MV increased with both valence (positive estimate) and arousal (positive estimate) (Fig. 2, lower right). For %RC, a significant arousal effect was found, indicating that with increasing arousal thoracic breathing tended to augment (positive estimate). SCL was higher for high-arousal excerpts than low-arousal excerpts (positive estimate). For Ti/Ttot, Vi, and HR neither the V–A-interaction model nor the V–A-only model indicated a significant covariation with valence or arousal ratings. For these variables, the V–A-only model is reported.

4. Discussion

In this study, we investigated the relationships between physiological parameters and subjective reports of affective valence and arousal while listening to environmental noises and musical passages.

Both similarities and differences emerge when comparing the results for noises and music. Increasing arousal was associated with faster breathing and higher MV for both types of stimuli. For music, this relation was present over the entire affective space, for noises essentially for positive stimuli, only. Increases in rated valence were associated to faster breathing and increases in MV for music but to slower breathing and decreases in MV for relatively low-arousal noises (Fig. 2). Differences were also obtained for SCL and HR. For noises, SCL was unrelated to the affective judgments, whereas HR augmented with increasing arousal. In contrast, for music HR was unrelated to reports of affective dimensions whereas SCL increased with rated arousal.

4.1. Affective judgments

For most subjects, the affective judgments varied widely across the two-dimensional space both for noises and musical fragments. However, for noises the negatively valent low-arousal area was relatively scantily occupied, whereas the negatively valent high-arousal area was strongly represented. This distribution of ratings is very similar to that reported by Bradley and Lang (2000) using a larger set of acoustic stimuli and also to that obtained previously for pictorial stimuli (Bradley and Lang, 2000).

Several studies have shown music to be a valid instrument to induce a wide variety of affective reactions (e.g. Rigg, 1964; Krumhansl, 1997; Nyklícek et al., 1997). In line with this, we were able to induce a broad range of affective responses including negative low-arousal ones (Witvliet, 1998). Furthermore, the musical stimuli were characterized by a relatively high consistency across participants in the arousal ratings. The valence ratings, on the contrary, were much more divergent with some fragments being rated by large subgroups of participants both positively and negatively. This is somehow in disagreement with those researchers claiming that emotional judgments of music are quite consistent across individuals (Krumhansl, 1997). Our finding might simply be due to the specific selection possibly including musical fragments that are 'emotionally ambiguous' and to which subjects may have associated differently emotionally charged events. Differences in musical preferences may in part be also responsible for this inconsistency, although no subject reported to dislike classical music, to which the majority of stimuli belonged. Alternatively, it might be that some subjects rated the samples as they actually felt and others rated how much they liked the music excerpts. The relationship between liking and underlying emotional state is not completely straightforward. For example, although a music passage may make us feel sad, at the same time we might be captured by its beauty and report this rather than our actual feeling of sadness (Schubert, 1996). Musical training is unlikely to have had any influence, because participants were very similar in this regard.

4.2. Relationships between affective judgments and physiological responses

4.2.1. Respiration

Increases in arousal were associated with faster breathing and higher Vi/Ti and MV for both types of stimuli. These findings confirm and extend previous reports on respiratory rate and flows (Nakamura, 1984; Boiten et al., 1994; Nyklícek et al., 1997; Boiten, 1998; Van Diest et al., 2001; Gomez et al., 2004a). More arousing music was also accompanied by enhanced thoracic breathing as previously obtained with pictures (Gomez et al., 2004a) and imagery (Rehwoldt, 1911).

Results also indicated a valence influence for both types of stimuli, which, however, was in opposite directions. In our study with pictures (Gomez et al., 2004a), we also found valence effects, that however, differ both from those for noises and for music. Thus, contrary to arousal, the influence of valence appears to be particularly stimulus-dependent. The dynamic nature of noises and in particular the rhythmical characteristics of music in contrast to the static nature of pictures may partly explain these differences. Emotions in music are conveyed through the structure of the music itself. Tempo and various rhythmical aspects are important determinants of affective responses to music (see Gabrielsson and Juslin, 2003 for review) and respiration tends to synchronize to musical rhythms (Harrer and Harrer, 1968; Kneutgen, 1970; Haas et al., 1986). The strong relations found between affective judgments and time parameters very likely reflect the synchronizing effect of music on respiration. A preliminary analysis on the relations between the technical characteristics of the excerpts and the psychophysiological measures indicates that arousing excerpts had faster tempo than calming ones and that positive excerpts had faster tempo and more dance-like rhythms than negative excerpts (Gomez and Danuser, 2004). Other things being equal, increased speed tends to be associated with positive emotions (Wedin, 1972; Dalla Bella et al., 2001). This is likely to explain why breathing accelerated and flows augmented with increasing pleasantness. Nakamura (1984) and Nyklícek et al. (1997) studies partly support the acceleration of breathing with increasing pleasantness observed here.

The synchronizing effect of music likely explains also why for music much stronger relationships between respiratory variables and affective ratings were obtained than for environmental noises. Even though some noises have a certain rhythm, this is relatively weak in comparison to music and consequently synchronization of respiration might not be as strong as with music. Alternatively, it could be that the selected set of noises did not include stimuli that belong to affectively potent categories (e.g. crying, laughing, erotica). However, Bradley and Lang (2000) failed to find strong effects using a much larger sample that included such groups of sounds.

Both for noises and music excerpts, no relationship between Vi and reports of valence and arousal was found. This finding parallels results from recent studies using perceptual stimuli (Boiten, 1998; Gomez et al., 2004a,b). It appears, that in relatively passive perception contexts, Vi is unaffected by variations in the affective content. Rafferty and Gardner (1996) have suggested that in contrast to the timing components, Vi is controlled dominantly by the chemical drive, whereas non-metabolic influences are weaker.

4.2.2. Skin conductance and heart rate

High-arousal music excerpts prompted higher SCL than low-arousal music excerpts. This finding is in accordance with reports of several studies (e.g. Shrift, 1957; Zimny and Weidenfeller, 1963; Nakamura, 1984; Witvliet, 1998). On the contrary, for noises no significant correlation between SCL and arousal ratings was observed. This cannot be due to a lack of variation in arousal levels, because ratings varied widely along the arousal dimension. High-arousal noises prompted higher HR than low-arousal noises whereas such effect was not found for music. Thus, the results for SCL and HR for the two types of acoustic stimuli were very different.

The findings for SCL and HR of the noise stimuli parallel those generally reported in imagery studies. SCL is unrelated to the emotional content of imagery (Lang et al., 1983; Vrana, 1993, 1994) and specifically to reports of valence and arousal (Witvliet and Vrana, 1995), whereas HR is higher during high-arousal than low-arousal imagery (Acosta and Vila, 1990; Witvliet and Vrana, 1995; Van Diest et al., 2001). One might speculate that the differences in SCL and HR between noises and music may be partly due to a difference in the stimulation of imagery. Hearing a specific noise may automatically evoke images of the respective object or event with the subject possibly imagining him- or herself as part of the scene, similarly to imagery paradigms. Although music can stimulate imagery, we often listen to it without generating mental images (Blood and Zatorre, 2001). In a follow-up study, we asked 16 students to report if they had visualized something while listening to five noises and five music excerpts randomly selected from the entire sample. On average, they answered affirmatively for 4.1 noises and for 1.1 music excerpts only. It is worth noting, that 20 subjects of our study reported to have generally tried to recognize the name of the composition from which the musical fragments were taken or the name of the composer, a rather intellectual activity likely to have hindered any sort of imagery and possibly reduced the emotional involvement.

Although it seems intuitively right that sedative music should slow down HR, and that stimulating music should speed it up, several studies have failed to find any relation between HR and indices of arousal (e.g. Ellis and Brighouse, 1952; Zimny and Weidenfeller, 1963; Burns et al., 1999; see Bartlett, 1996 for review), whereas others suggest that music can have strong effects on HR (Krumhansl, 1997; Nyklícek et al., 1997; Witvliet, 1998; Iwanaga and Moroki, 1999). In particular, Witvliet (1998) found HR to be significantly higher during high-arousal music, compared to low-arousal music. Several factors may account for this inconsistency (e.g. task instructions to the subjects, stimulus intensity, subject attitude, duration of music, see Dainow, 1977). In Witvliet's study (1998) subjects were asked to try to experience the emotion expressed in the music. According to Witvliet (1998), this instruction is similar to the task of personally experiencing affective situations in imagery paradigms and might accentuate HR responsivity to the affective arousal of the stimuli. Further, keeping one's eyes closed while listening to music as in the studies by Nyklícek et al. (1997), Witvliet (1998) and Iwanaga and Moroki (1999) may also favor emotional engagement possibly by encouraging imagery. On the contrary, asking to merely listen to the music as in the studies by Ellis and Brighouse (1952) and Zimny and Weidenfeller (1963) does not emphasize the task as an emotional one.

In conclusion, our study has shown that breathing responses to environmental noises and music excerpts are organized to a certain degree along the affective dimensions of valence and arousal. Although both noises and music are sound-vibrations, differences in the relations between affective judgments and physiological responses were observed suggesting differences in the processing of the two types of acoustic stimuli. We hypothesize that one important difference may be the degree to which music and noises stimulated imagery in the participants.

Acknowledgments

The authors thank Ronny Ott for his assistance in conducting this experiment and Dr Werner A. Stahel for his statistical advise.

References

Acosta, A., Vila, J., 1990. Emotional imagery: effect of autonomic response information on physiological arousal. Cognition Emotion 4, 145–160.

Bartlett, D.L., 1996. Physiological responses to music and sound stimuli. In: Hodges, D.A. (Ed.), Handbook of Music Psychology. IMR Press, San Antonio, pp. 343–385.

Blood, A.J., Zatorre, R.J., 2001. Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. Proc. Natl. Acad. Sci. USA 98 (20), 11 818–11 823.

Boiten, F.A., 1998. The effects of emotional behaviour on components of the respiratory cycle. Biol. Psychol. 49, 29–51.

Boiten, F.A., Frijda, N.H., Wientjes, C.J.E., 1994. Emotions and respiratory patterns: review and critical analysis. Int. J. Psychophysiol. 17, 103–128.

Bradley, M.M., Lang, P.J., 1994. Measuring emotion: the self-assessment manikin and the semantic differential. J. Behav. Ther. Exp. Psychiatry 25, 49–59.

Bradley, M.M., Lang, P.J., 2000. Affective reactions to acoustic stimuli. Psychophysiology 37, 204–215.

Burns, J., Labbé, E., Williams, K., McCall, J., 1999. Perceived and physiological indicators of relaxation: as different as Mozart and Alice in chains. Appl. Psychophysiol. Biofeedback 24, 197–202.

Dainow, E., 1977. Physical effects and motor responses to music. J. Res. Music Educ. 25, 211–221.

Dalla Bella, S., Peretz, I., Rousseau, L., Gosselin, N., Ayotte, J., Lavoie, A., 2001. Development of the happy–sad distinction in music appreciation – Does tempo emerge earlier than mode? In: Zatorre, R.J., Peretz, I. (Eds.), The Biological Foundations of Music. The New York Academy of Sciences, NY, pp. 436–438.

Dibben, N., 2001. What do we hear, when we hear music?: music perception and musical material. Musicae Scientiae 2, 161–194.

Ellis, D.S., Brighouse, G., 1952. Effects of music on respiration and heart-rate. Am. J. Psychol. 65, 39–47.

Fridlund, A.J., 1991. Sociality of solitary smiling: potentiation by an implicit audience. J. Pers. Soc. Psychol. 60, 229–240.

Gabrielsson, A., Juslin, P.N., 2003. Emotional expression in music. In: Davidson, R.J., Scherer, K.R., Hill Goldsmith, H. (Eds.), Handbook of Affective Sciences. Oxford university press, New York, pp. 503–534.

Gaver, W.W., 1993. What in the world do we hear? An ecological approach to auditory event perception. Ecol. Psychol. 5, 1–29.

Gomez, P., Danuser, B, 2004. Relationship between musical features and psychophysiological measures of emotion. Manuscript in preparation.

Gomez, P., Stahel, W.A., Danuser, B, 2004a. Respiratory responses during affective picture viewing. Biol. Psychol., article in press.

Gomez, P., Zimmermann, P., Guttormsen-Schär, S., Danuser, B. 2004b. Respiratory responses associated with affective processing of film stimuli. Biol. Psychol., under revision.

Haas, F., Distenfeld, S., Axen, K., 1986. Effects of perceived musical rhythm on respiratory pattern. J. Appl. Physiol. 61, 1185–1191.

Hamm, A.O., Schupp, H.T., Weike, A.I., 2003. Motivational organization of emotions: autonomic changes, cortical responses, and reflex modulation. In: Davidson, R.J., Scherer, K.R., Hill Goldsmith, H. (Eds.), Handbook of Affective Sciences. Oxford university press, New York, pp. 187–211.

Harrer, G., Harrer, H., 1968. Musik, emotion und vegetativum. Wiener Med. Wochenschr. 118, 966–971.

Hedeker, D., Gibbons, R.D., 1996. MIXREG: a computer program for mixed-effects regression analysis with autocorrelated errors. Comput. Methods Prog. Biomed. 49, 229–252.

Iwanaga, M., Moroki, Y., 1999. Subjective and physiological responses to music stimuli controlled over activity and preference. J. Music Therapy 33, 26–38.

Jäncke, L., 1994. An EMG investigation of the covariation of facial muscles during the presentation of affect-laden stimuli. J. Psychophysiol. 8, 1–10.

Kneutgen, J., 1970. Eine Musikform und ihre biologische Funktion. Über die Wirkungsweise der Wiegenlieder. Zeitschrift für experimentelle angewandte Psychologie 17, 245–265.

Krumhansl, C.L., 1997. An exploratory study of musical emotions and psychophysiology. Can. J. Exp. Psychol. 51, 336–352.

Lang, P.J., Levin, D.N., Miller, G.A., Kozak, M.J., 1983. Fear behavior, fear imagery, and the psychophysiology of emotion: the problem of affective response integration. J. Abnormal Psychol. 92, 276–306.

Nakamura, H., 1984. Effects of musical emotionality upon GSR and respiration rate: the relationship between verbal reports and physiological responses. Shinrigaku Kenkyu 55, 47–50.

Nyklícek, I., Thayer, J.F., Van Doornen, L.J.P., 1997. Cardiorespiratory differentiation of musically-induced emotions. J. Psychophysiol. 11, 304–321.

Rafferty, G.F., Gardner, W.N., 1996. Control of the respiratory cycle in conscious humans. J. Appl. Physiol. 81, 1744–1753.

Rehwoldt, F., 1911. Über respiratorische Affektsymptome. Psychologische Studien 7, 141–195.

Rigg, M.G., 1964. The mood effects of music: a comparison of data from four investigators. J. Psychol. 58, 427–438.

Ritz, T., Alatupa, S., Thöns, M., Dahme, B., 2002. Effects of affective picture viewing and imagery on respiratory resistance in non-asthmatic individuals. Psychophysiology 39, 86–94.

Sackner, M.A., Watson, H., Belsito, A.S., Feinerman, D., Suarez, M., Gonzalez, G. et al., 1989. Calibration of respiratory inductive plethysmography during natural breathing. J. Appl. Physiol. 66, 410–420.

Schubert, E., 1996. Enjoyment of negative emotions in music: an associative network explanation. Psychol. Music 24, 18–28.

Shrift, D.C., 1957. The galvanic skin response to two contrasting types of music. In: Gaston, E.T. (Ed.), Music Therapy. Allen Press, Lawrence, KS, pp. 235–239.

Van Diest, I., Winters, W., Devriese, S., Vercamst, E., Han, J.N., Van de Woestijne, K.P. et al., 2001. Hyperventilation beyond fight/flight: respiratory responses during emotional imagery. Psychophysiology 38, 961–968.

Witvliet, C.V.O. 1998. The impact of music-prompted emotional valence and arousal on self-report, autonomic, facial EMG, and startle responses across experimental contexts. Doctoral Thesis, Purdue University, UMI number: 9819062.

Witvliet, C.V.O., Vrana, S.R., 1995. Psychophysiological responses as indices of affective dimensions. Psychophysiology 32, 436–443.

Vrana, S.R., 1993. The psychophysiology of disgust: differentiating negative emotional contexts with facial EMG. Psychophysiology 30, 279–286.

Vrana, S.R., 1994. Startle reflex response during sensory modality specific disgust, anger, and neutral imagery. J. Psychophysiol. 8, 211–218.

Wedin, L., 1972. A multidimensional study of perceptual-emotional qualities in music. Scand. J. Psychol. 13, 241–257.

Wientjes, C.J.E., 1992. Respiration in psychophysiology: methods and applications. Biol. Psychol. 34, 179–203.

Zimny, G.H., Weidenfeller, E.W., 1963. Effects of music upon GSR and heart-rate. Am. J. Psychol. 76, 311–314.