Subjective Emotional Responses to Musical Structure, Expression and Timbre Features: A Synthetic Approach

Sylvain Le Groux[1], Paul F.M.J. Verschure[1,2]

[1]SPECS, Universitat Pompeu Fabra
[2]ICREA, Barcelona
{sylvain.legroux, paul.verschure}@upf.edu

Abstract. Music appears to deeply affect emotional, cerebral and physiological states, and its effect on stress and anxiety has been established using a variety of self-report, physiological, and observational means. Yet, the relationship between specific musical parameters and emotional responses is still not clear. One issue is that precise, replicable and independent control of musical parameters is often difficult to obtain from human performers. However, it is now possible to generate expressive musical material such as pitch, velocity, articulation, tempo, scale, mode, harmony and timbre using synthetic music systems. In this study, we use a synthetic music system called the SMuSe, to generate a set of well-controlled musical stimuli, and analyze the influence of musical structure, performance variations and timbre on emotional responses.The subjective emotional responses we obtained from a group of 13 participants on the scale of valence, arousal and dominance were similar to previous studies that used human-produced musical excerpts. This validates the use of a synthetic music system to evoke and study emotional responses in a controlled manner.

Keywords: music-evoked emotion, synthetic music system

1 Introduction

It is widely acknowledged that music can evoke emotions and synchronized reactions of experiential, expressive and physiological components of emotion have been observed while listening to music [1]. A key question is how musical parameters can be mapped to emotional states of valence, arousal and dominance. In most of the cases, studies on music and emotion are based on the same paradigm: one measures emotional responses while the participant is presented with an excerpt of recorded music. These recordings are often extracted from well-known pieces of the repertoire and interpreted by human performers who follow specific expressive instructions. One drawback of this methodology is that expressive interpretation can vary quite a lot from one performer to another, which compromises the generality of the results. Moreover, it is difficult, even

9th International Symposium on Computer Music Modelling and Retrieval (CMMR 2012)
19-22 June 2012, Queen Mary University of London

for a professional musician, to accurately modulate one single expressive dimension independently of the others. Many dimensions of the stimuli might not be controlled for. Besides, pre-made recordings do not provide any control over the musical content and structure.

In this paper, we propose to tackle these limitations by using a synthetic composition system called the SMuSe [2,3] to generate stimuli for the experiment. The SMuSe allows to generate synthetic musical pieces and to modulate expressive musical material such as pitch, velocity, articulation, tempo, scale, mode, harmony and timbre. It provides accurate, replicable and independent control over perceptually relevant time-varying dimensions of music.

Emotional responses to music most probably involve different types of mechanisms such as cognitive appraisal, brain stem reflexes, contagion, conditioning, episodic memory, or expectancy [4]. In this study, we focused on the direct relationship between basic perceptual acoustic properties and emotional responses of a reflexive type. As a first approach to assess the participants' emotional responses, we looked at their subjective responses following the well-established three dimensional theory of emotions (valence, arousal and dominance) illustrated by the Self Assessment Manikin (SAM) scale [5,6].

2 Methods

2.1 Stimuli

This experiment investigates the effects of a set of well-defined musical parameters within the three main musical determinants of emotions, namely structure, performance and timbre. In order to obtain a well-parameterized set of stimuli, all the sound samples were synthetically generated. The composition engine SMuSe[1] allowed the modulation of macro-level musical parameters (contributing to structure, expressivity) via a graphical user interface [2,3], while the physically-informed synthesizer PhySynth[2] allowed to control micro-level sound parameters [7] (contributing to timbre). Each parameter was considered at three different levels (Low, Medium, High). All the sound samples[3] were 5 s. long and normalized in amplitude with the Peak Pro[4] audio editing and processing software. .

Musical Structure: To look at the influence of musical structure on emotion, we focused on two simple but fundamental structural parameters namely register (Bass, Tenor and Soprano) and mode (Random, C Minor, C Major). A total of 9 sound samples (3 Register * 3 Mode levels) were generated by SMuSe (Figure 1).

[1] http://goo.gl/Vz1ti
[2] http://goo.gl/zRLuC
[3] http://goo.gl/5iRM0
[4] http://www.bias-inc.com/

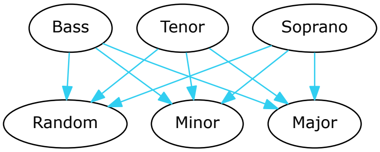

Fig. 1. Musical structure samples: Register and Mode are modulated over 9 sequences (3*3 combinations)

Expressivity Parameters: Our study of the influence of musical performance parameters on emotion relies on three expressive parameters, namely tempo, dynamics, and articulation that are commonly modulated by live musicians during performance. A total of 27 sound samples (3 Tempo * 3 Dynamics * 3 Articulation) were generated by SMuSe (Figure 2).

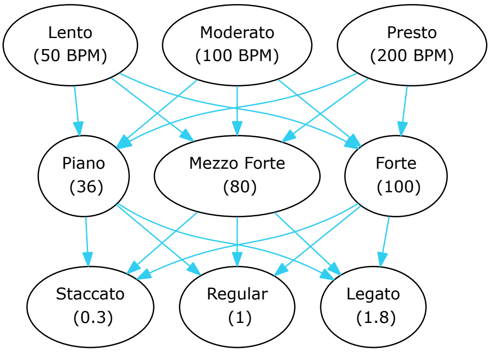

Fig. 2. Musical performance samples: 3 performance parameters were modulated over 27 musical sequences (3*3*3 combinations of Tempo (BPM), Dynamics (MIDI velocity value) and Articulation (duration multiplication factor) levels).

Timbre: For timbre, we focused on parameters that relate to the three main dimension of timbre namely brightness (controlled by tristimulus value), attack-time and spectral flux (controlled by damping). A total of 27 sound samples (3 Attack Time * 3 Brightness * 3 Damping) were generated by PhySynth (Figure 3). For a more detailed description of the timbre parameters, refer to [7].

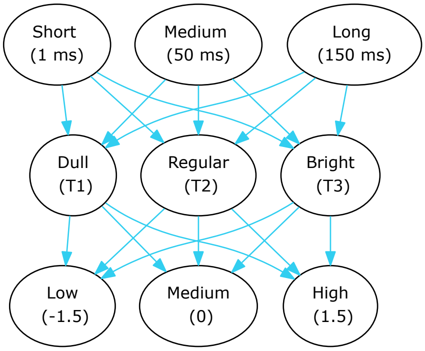

Fig. 3. Timbre samples: 3 timbre parameters are modulated over 27 samples (3*3*3 combinations of Attack (ms), Brightness (tristimulus band), Damping (relative damping α)). The other parameters of PhySynth were fixed: decay=300ms, sustain=900ms, release=500ms and global damping $\alpha_g = 0.23$.

2.2 Procedure

We investigated the influence of different sound features on the emotional state of the patients using a fully automated and computer-based stimulus presentation and response registration system. In our experiment, each subject was seated in front of a PC computer with a 15.4" LCD screen and interacted with custom-made stimulus delivery and data acquisition software called PsyMuse[5] (Figure 4) made with the Max-MSP [6] programming language [8]. Sound stimuli were presented through headphones (K-66 from AKG).

At the beginning of the experiment, the subject was exposed to a sinusoidal sound generator to calibrate the sound level to a comfortable level and was explained how to use PsyMuse's interface (Figure 4). Subsequently, a number of sound samples with specific sonic characteristics were presented together with the different scales (Figure 4) in three experimental blocks (structure, performance, timbre) containing all the sound conditions presented randomly.

For each block, after each sound, the participants rated the sound in terms of its emotional content (valence, arousal, dominance) by clicking on the SAM manikin representing her emotion [6]. The participants were given the possibility to repeat the playback of the samples. The SAM 5 points graphical scale gave a score (from 0 to 4) where 0 corresponds to the most dominated, aroused and positive and 4 to the most dominant, calm and negative (Figure 4). The data was automatically stored into a SQLite[7] database composed of a table for

[5] http://goo.gl/fxOOL
[6] http://cycling74.com/
[7] http://www.sqlite.org/

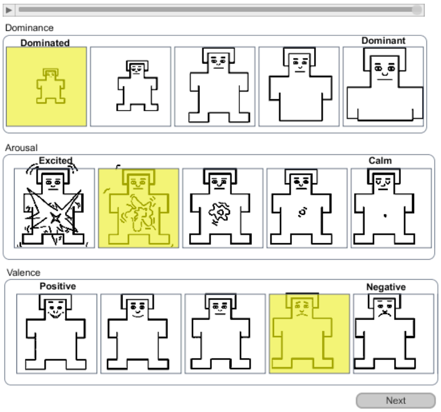

Fig. 4. The presentation software PsyMuse uses the SAM scales (axes of Dominance, Arousal and Valence) [6] to measure the participant's emotional responses to a database of sounds.

demographics and a table containing the emotional ratings. SPSS[8] (from IBM) statistical software suite was used to assess the significance of the influence of sound parameters on the affective responses of the subjects .

2.3 Participants

A total of N=13 university students (5 women, $M_{age} = 25.8$, range=22-31) with normal hearing took part in the pilot experiment. The experiment was conducted in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki[9]. Six of the subjects had musical background ranging from two to seven years of instrumental practice.

3 Results

The experiment followed a blocked within-subject design where for each of the three block (structure, performance, timbre) every participant experienced all the conditions in random order.

3.1 Musical Structure

To study the emotional effect of the structural aspects of music, we looked at two independent factors (register and mode) with three levels each (soprano, bass, tenor and major, minor, random respectively) and three dependent variables (Arousal, Valence, Dominance). The Kolmogorov-Smirnov test showed that the

[8] http://www.spss.com/
[9] http://www.wma.net/en/30publications/10policies/b3/index.html

data is normally distributed. Hence, we carried a Two-Way Repeated Measure Multivariate Analysis of Variance (MANOVA).

The analysis showed a multivariate effect for the $mode * register$ interaction $V(12, 144) = 1.92, p < 0.05$. Mauchly tests indicated that assumption of sphericity is met for the main effects of register and mode as well as for the interaction effect. Hence we did not correct the F-ratios for follow-up univariate analysis.

Follow-up univariate analysis revealed an effect of **register** on **arousal** $F(2, 24) = 2.70, p < 0.05$ and **mode** on **valence** $F(2, 24) = 3.08, p < 0.05$ as well as a **mode** $*$ **register** interaction effect on arousal $F(4, 48) = 2.24, p < 0.05$, dominance $F(4, 48) = 2.64, p < 0.05$ and valence $F(4, 48) = 2.73, p < 0.05$ (Cf. Table 1).

	ANOVAs		
	Register	**Mode**	**Register * Mode**
Arousal	F(2,24)=2.70, *p<.05	NS	F(4,48)=2.238, *p<0.05
Valence	NS	F(2, 24)=3.079, *p<0.05	F(4,48)=2.636, *p<0.05
Dominance	NS	NS	F(4,48)=2.731, *p<0.05

Table 1. Effect of mode and register on the emotional scales of arousal, valence and dominance: statistically significant effects.

A post-hoc pairwise comparison with Bonferroni correction showed a significant mean difference of -0.3 between High and Low register and of -0.18 between High and Medium on the arousal scale (Figure 5 B). High register appeared more arousing than medium and low register.

A pairwise comparison with Bonferroni correction showed a significant mean difference of -0.436 between random and major (Figure 5 A). Random mode was perceived as more negative than major mode.

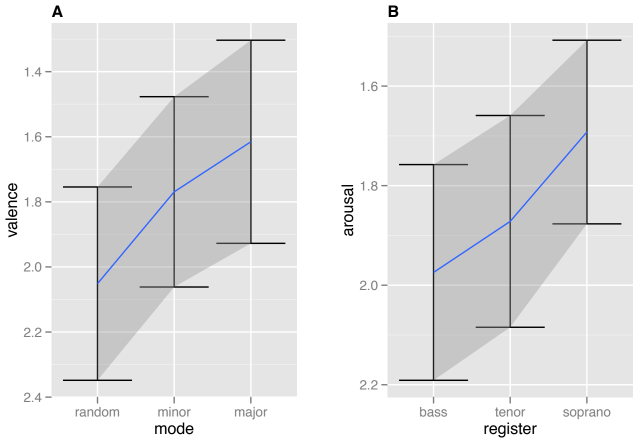

Fig. 5. Influence of structural parameters (register and mode) on arousal and valence. A) A musical sequence played using random notes and using a minor scale is perceived as significantly more negative than a sequence played using a major scale. **B)** A musical sequence played in the soprano range (respectively bass range) is significantly more (respectively less) arousing than the same sequence played in the tenor range. Estimated Marginal Means are obtained by taking the average of the means for a given condition.

The interaction effect between mode and register suggests that the random mode has a tendency to make a melody with medium register less arousing (Figure 6, A). Moreover, the minor mode tended to make high register more positive and low register more negative (Figure 6, B). The combination of high register and random mode created a sensation of dominance (Figure 6, C).

3.2 Expressive Performance Parameters

To study the emotional effect of some expressive aspects of music during performance, we decided to look at three independent factors (Articulation, Tempo, Dynamics) with three levels each (high, low, medium) and three dependent variables (Arousal, Valence, Dominance). The Kolmogorov-Smirnov test showed that the data was normally distributed. We did a Three-Way Repeated Measure Multivariate Analysis of Variance.

The analysis showed a multivariate effect for **Articulation** $V(4.16, 3) < 0.05$, **Tempo** $V(11.6, 3) < 0.01$ and **dynamics** $V(34.9, 3) < 0.01$. No interaction effects were found.

Mauchly tests indicated that the assumption of sphericity was met for the main effects of articulation, tempo and dynamics on arousal and valence but not dominance. Hence we corrected the F-ratios for univariate analysis for dominance with Greenhouse-Geisser.

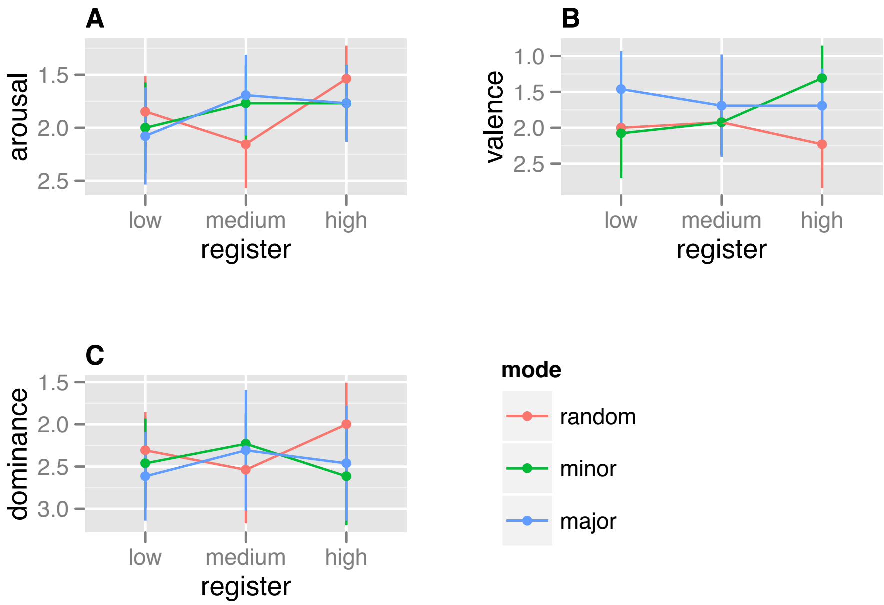

Fig. 6. Structure: interaction between mode and register for arousal, valence and dominance. A) When using a random scale, a sequence in the tenor range (level 3) becomes less arousing **B)** When using a minor scale , a sequence played within the soprano range becomes the most positive. **C)** When using a random scale, bass and soprano sequences are the most dominant whereas tenor becomes the less dominant.

	ANOVAs		
	Articulation	**Tempo**	**Dynamics**
Arousal	F(2,24)=6.77, **p<0.01	F(2,24)=27.1, ***p<0.001	F(2,24)=45.78, ***p<0.001
Valence	F(2,24)=7.32, **p<0.01	F(2, 24)=4.4, *p<0.05	F(2,24)=19, ***p<0.001
Dominance	NS	F(1.29,17.66)=8.08, **p<0.01	F(2,24)=9.7, **p<0.01

Table 2. Effect of articulation, tempo and dynamics on self-reported emotional responses on the scale of valence, arousal and dominance**:** statistically significant effects.

Arousal Follow-up univariate analysis revealed an effect of **articulation** $F(6.76, 2) < 0.01$, **tempo** $F(27.1, 2) < 0.01$, and **dynamics** $F(45.77, 2) < 0.05$ on arousal (Table 2).

A post-hoc pairwise comparison with Bonferroni correction showed a significant mean difference of 0.32 between the **articulation** staccato and legato (Figure 7 A). The musical sequence played staccato was perceived as more arousing.

A pairwise comparison with Bonferroni correction showed a significant mean difference of -1.316 between high **tempo** and low tempo and -0.89 between high and medium tempo (Figure 7 B). This shows that a musical sequence with higher tempi was perceived as more arousing.

A pairwise comparison with Bonferroni correction showed a significant mean difference of -0.8 between forte and piano **dynamics**, -0.385 between forte and regular and 0.41 between piano and regular (Figure 7 C). This shows that a musical sequence played at higher dynamics was perceived as more arousing.

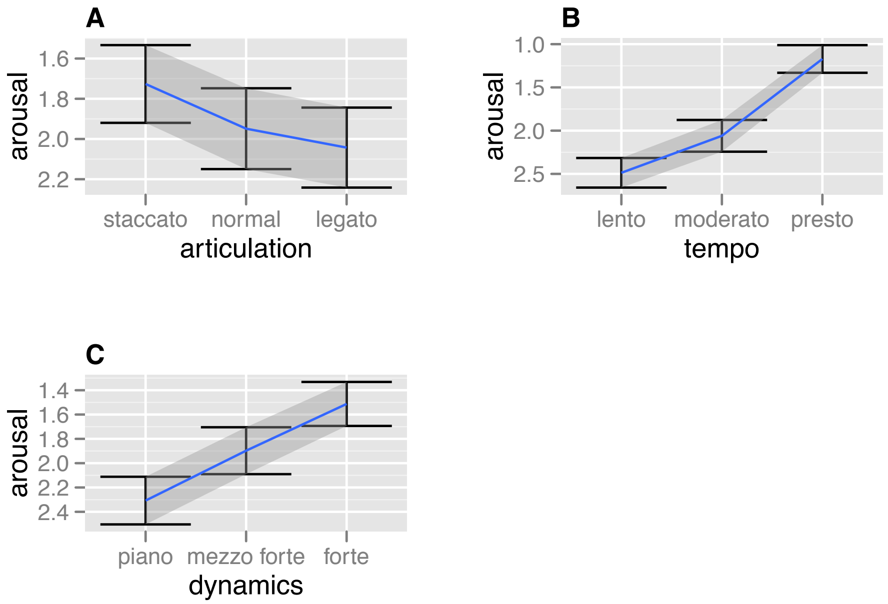

Fig. 7. Effect of performance parameters (Articulation, Tempo and Dynamics) on Arousal. A) A sequence played with articulation staccato is more arousing than legato **B)** A sequence played with the tempo indication presto is more arousing than both moderato and lento. **C)** A sequence played forte (respectively piano) was more arousing (respectively less arousing) than the same sequence played mezzo forte.

Valence Follow-up univariate analysis revealed an effect of **articulation** $F(7.31, 2) < 0.01$, **tempo** $F(4.3, 2) < 0.01$, and **dynamics** $F(18.9, 2) < 0.01$ on valence (Table 2)

A post-hoc pairwise comparison with Bonferroni correction showed a significant mean difference of -0.32 between the **articulation** staccato and legato (Figure 7 A). The musical sequences played with shorter articulations were perceived as more positive.

A pairwise comparison with Bonferroni correction showed a significant mean difference of 0.48 between high **tempo** and medium tempo (Figure 8 B). This shows that sequences with higher tempi tended be perceived as more negatively valenced.

A pairwise comparison with Bonferroni correction showed a significant mean difference of 0.77 between high and low **dynamics** and -0.513 between low and medium. (Figure 8 C). This shows that musical sequences played with higher dynamics were perceived more negatively.

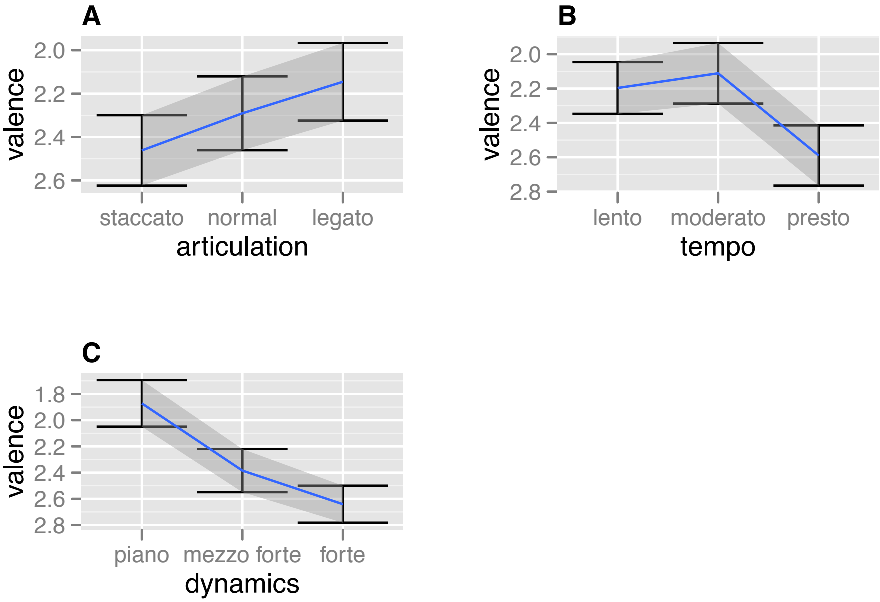

Fig. 8. Effect of performance parameters (Articulation, Tempo and Dynamics) on Valence. A) A musical sequence played staccato induce a more negative reaction than when played legato **B)** A musical sequence played presto is also inducing a more negative response than played moderato. **C)** A musical sequence played forte (respectively piano) is rated as more negative (respectively positive) than a sequence played mezzo forte.

Dominance Follow-up univariate analysis revealed an effect **Tempo** $F(8, 2) < 0.01$, and **dynamics** $F(9.7, 2) < 0.01$ on valence (Table 2).

A pairwise comparison with Bonferroni correction showed a significant mean difference of -0.821 between high **tempo** and low tempo and -0.53 between high tempo and medium tempo (Figure 9 A). This shows that sequences with higher tempi tended to make the listener feel dominated.

A pairwise comparison with Bonferroni correction showed a significant mean difference of -0.55 between high and low **dynamics** and 0.308 between low and medium (Figure 9 B). This shows that when listening to musical sequences played with higher dynamics, the participants felt more dominated.

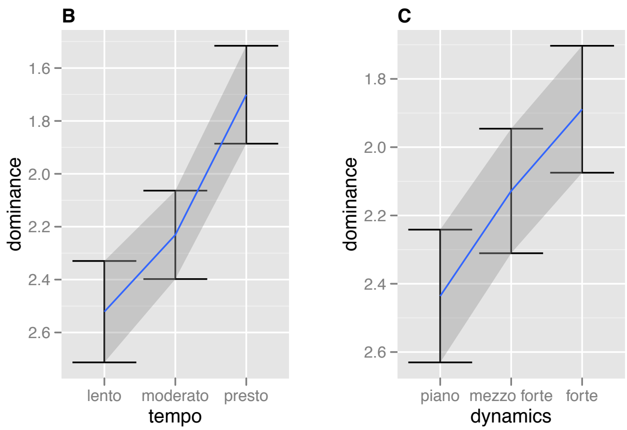

Fig. 9. Effect of performance parameters (Tempo and Dynamics) on Dominance. A) A musical sequence played with a tempo presto (repectively lento) is considered more dominant (respectively less dominant) than played moderato **B)** A musical sequence played forte (respectively piano) is considered more dominant (respectively less dominant) than played mezzo-forte

3.3 Timbre

To study the emotional effect of the timbral aspects of music, we decided to look at three independent factors known to contribute to the perception of Timbre [9,10,11] (Attack time, Damping and Brightness) with three levels each (high, low, medium) and three dependent variables (Arousal, Valence, Dominance). The Kolmogorov-Smirnov test showed that the data is normally distributed. We did a Three-Way Repeated Measure Multivariate Analysis of Variance.

The analysis showed a multivariate effect for **brightness** $V(6, 34) = 3.76, p < 0.01$, **damping** $V(6, 34) = 3.22, p < 0.05$ and **attack time** $V(6, 34) = 4.19, p < 0.01$ and an interaction effect of **brightness** $*$ **damping** $V(12, 108) = 2.8 < 0.01$

Mauchly tests indicated that assumption of sphericity was met for the main effects of articulation, tempo and dynamics on arousal and valence but not dominance. Hence we corrected the F-ratios for univariate analysis for dominance with Greenhouse-Geisser.

	ANOVAs			
	Brightness	**Damping**	**Attack**	**Brightness* Damping**
Arousal	F(2,18)=29.09, ***p<0.001	F(2,18)=16.03, ***p<0.001	F(2,18)=3.54, *p<0.05	F(4,36)=7.47, ***p<0.001
Valence	F(2,18)=5.99, **p<0.01	NS	F(2,18)=7.26, **p<0.01	F(4,36)=5.82, **p<0.01
Dominance	F(1.49,13.45) =6.55, *p<0.05	F(1.05,10.915) =4.7, *p<0.05	NS	NS

Table 3. Effect of brightness, damping and attack on self-reported emotion on the scales of valence, arousal and dominance: statistically significant effects.

Arousal Follow-up univariate analysis revealed the main effects of **Brightness** $F(2,18) = 29.09 < 0.001$, **Damping** $F(2,18) = 16.03 < 0.001$, **Attack** $F(2,18) = 3.54 < 0.05$, and interaction effect **Brightness * Damping** $F(4,36) = 7.47, p < 0.001$ on Arousal (Figure 3).

A post-hoc pairwise comparison with Bonferroni correction showed a significant mean difference between high, low and medium **brightness**. There was a significant difference of -1.18 between high and low brightness, -0.450 between high and medium and -0.73 between medium and low. The brighter the sounds the more arousing.

Similarly significant mean difference of .780 between high and low **damping** and -0.37 between low and medium damping were found. The more damped, the less arousing.

For the **attack time** parameter, a significant mean difference of -0.11 was found between short and medium attack. Shorter attack time were found more arousing.

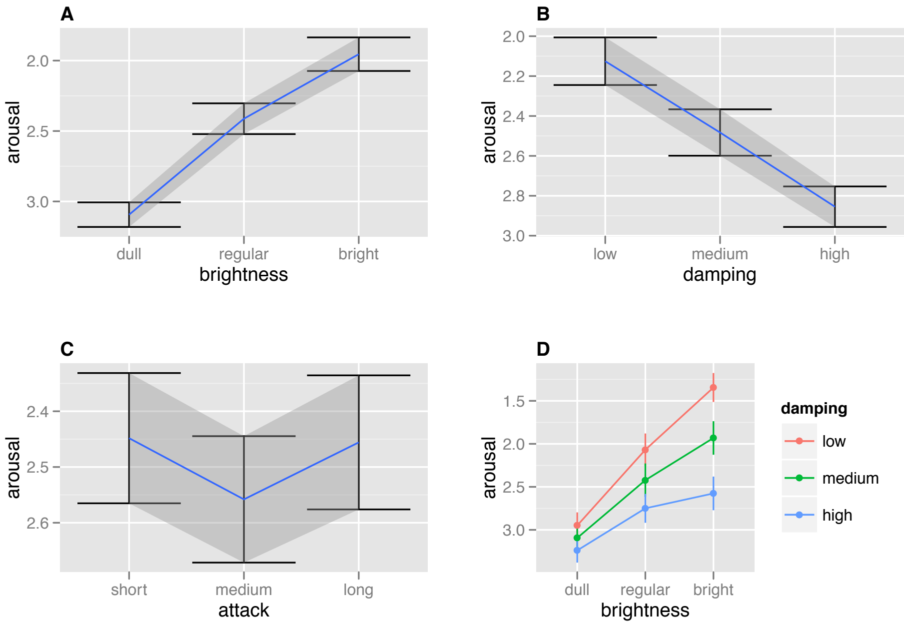

Fig. 10. Effect of timbre parameters (Brightness, Damping and Attack time) on Arousal. A) Brighter sounds induced more arousing responses. **B)** Sounds with more damping were less arousing. **C)** Sounds with short attack time were more arousing than medium attack time. **D)** Interaction effects show that less damping and more brightness lead to more arousal.

Valence Follow-up univariate analysis revealed main effects of **Brightness** $F(2, 18) = 5.99 < 0.01$ and **Attack** $F(2, 18) = 7.26 < 0.01$, and interaction effect **Brightness * Damping** $F(4, 36) = 5.82, p < 0.01$ on Valence (Figure 3).

Follow up pairwise comparisons with Bonferroni correction showed significant mean differences of 0.78 between high and low **brightness** and 0.19 between short and long **attacks** and long and medium attacks. Longer attacks and brighter sounds were perceived as more negative (Figure 11).

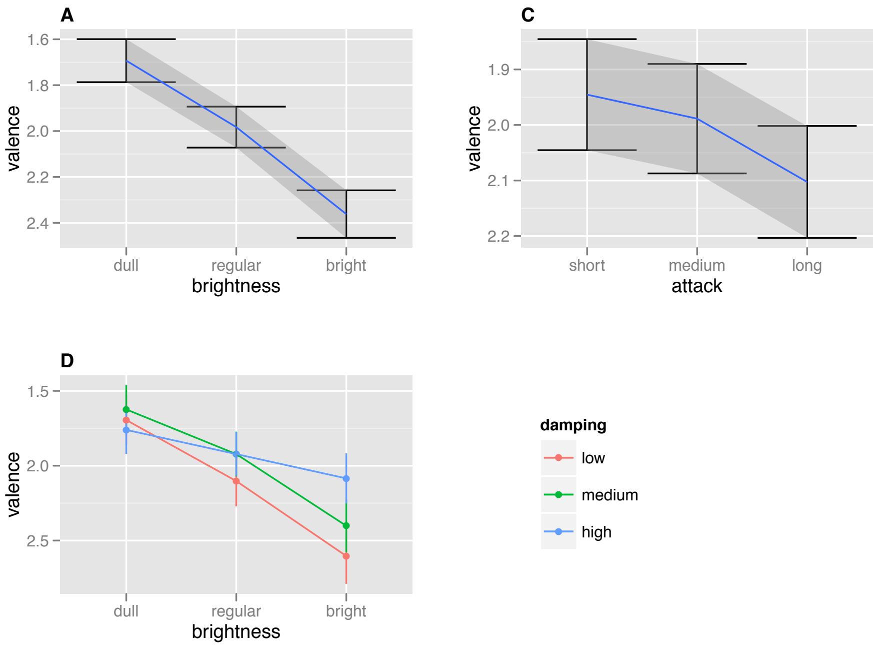

Fig. 11. Effect of timbre parameters (Brightness, Damping and Attack time) on Valence. A) Longer attack time are perceived as more negative **B)** Bright sounds tend to be perceived more negatively than dull sounds **C)** Interaction effects between damping and brightness show that a sound with high damping attenuates the negative valence due to high brightness.

Dominance Follow-up univariate analysis revealed main effects of **Brightness** $F(1.49, 13.45) = 6.55, p < 0.05$ and **Damping** $F(1.05, 10.915) = 4.7, p < 0.05$ on Dominance (Figure 3).

A significant mean difference of -0.743 was found between high and low **brightness**. The brighter the more dominant.

A significant mean difference of 0.33 was found between medium and low **damping** factor. The more damped the less dominant.

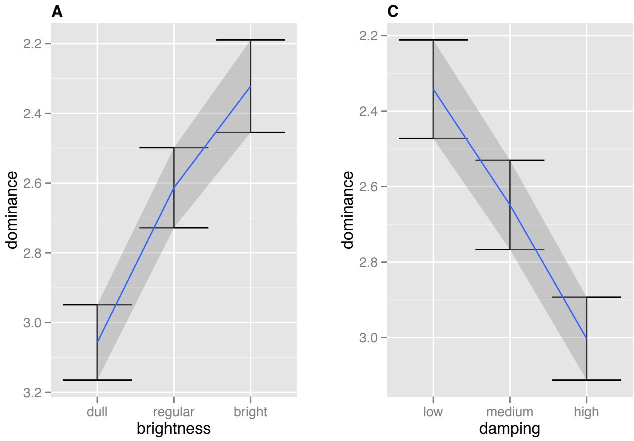

Fig. 12. Effect of timbre parameters (Brightness and Damping) on Dominance. A) Bright sounds are perceived as more dominant than dull sounds **B)** A sound with medium damping is perceived as less dominant than low damping.

4 Conclusions

This study validates the use of the SMuSe as an "affective music engine". The different levels of musical parameters that were experimentally tested evoked significantly different emotional responses. The tendency of minor mode to increase negative valence and of high register to increase arousal (Figure 5) corroborates the results of [12,13], and is complemented by interaction effects (Figure 6). The tendency of short articulation to be more arousing and more negative (Figure 7 and 8) confirms results reported in [14,15,16]. Similarly, higher tempi have a tendency to increase arousal and decrease valence (Figure 7 and 8) are also reported in [14,15,12,13,17,16]. The present study also indicates that higher tempi are perceived as more dominant (Figure 9). Musical sequences that were played louder were found more arousing and more negative (Figure 7 and 8) which is also reported in[14,15,12,13,17,16], but also more dominant (Figure 9). The fact that higher brightness tends to evoke more arousing and negative responses (Figure 10 and 11) has been reported (but in terms of number of harmonics in the spectrum) in [13]. Additionally, brighter sounds are perceived as more dominant (Figure 12). Damped sounds are less arousing and dominant (Figure 10 and 12). Sharp attacks are more arousing and more positive (Figure 10 and 11). Similar results were also reported by [14]. Additionally, this study revealed interesting interaction effects between damping and brightness (Figure 10 and 11).

Most of the studies that investigate the determinants of musical emotion use recordings of musical excerpts as stimuli. In this experiment, we looked at the effect of a well-controlled set of synthetic stimuli (generated by the SMuSe) on the listener's emotional responses. We developed an automated test procedure that assessed the correlation between a few parameters of musical structure, expressivity and timbre with the self-reported emotional state of the participants. Our results generally corroborated the results of previous meta-analyses [15], which suggests our synthetic system is able to evoke emotional reactions as well as "real" musical recordings. One advantage of such a system for experimental studies though, is that it allows for precise and independent control over the musical parameter space, which can be difficult to obtain, even from professional musicians. Moreover with this synthetic approach, we can precisely quantify the level of the specific musical parameters that led to emotional responses on the scale of arousal, valence and dominance. These results pave the way for an interactive approach to the study of musical emotion, with potential application to interactive sound-based therapies. In the future, a similar synthetic approach could be developed to further investigate the time-varying characteristics of emotional reactions using continuous two-dimensional scales and physiology [18,19].

References

1. L.-O. Lundqvist, F. Carlsson, P. Hilmersson, and P. N. Juslin, "Emotional responses to music: experience, expression, and physiology," *Psychology of Music* **37**(1), pp. 61–90, 2009.
2. S. Le Groux and P. F. M. J. Verschure, *Music Is All Around Us: A Situated Approach to Interactive Music Composition*. Exeter: Imprint Academic, April 2011.
3. S. Le Groux and P. F. M. J. Verschure, "Situated interactive music system: Connecting mind and body through musical interaction," in *Proceedings of the International Computer Music Conference*, Mc Gill University, (Montreal, Canada), August 2009.
4. P. N. Juslin and D. Västfjäll, "Emotional responses to music: the need to consider underlying mechanisms," *Behav Brain Sci* **31**, pp. 559–75; discussion 575–621, Oct 2008.
5. J. A. Russell, "A circumplex model of affect," *Journal of Personality and Social Psychology* **39**, pp. 345–356, 1980.
6. P. Lang, "Behavioral treatment and bio-behavioral assessment: computer applications," in *Technology in Mental Health Care Delivery Systems*, J. Sidowski, J. Johnson, and T. Williams, eds., pp. 119–137, 1980.
7. S. Le Groux and P. F. M. J. Verschure, "Emotional responses to the perceptual dimensions of timbre: A pilot study using physically inspired sound synthesis," in *Proceedings of the 7th International Symposium on Computer Music Modeling*, (Malaga, Spain), June 2010.
8. D. Zicarelli, "How I learned to love a program that does nothing," *Computer Music Journal* (26), pp. 44–51, 2002.
9. S. McAdams, S. Winsberg, S. Donnadieu, G. De Soete, and J. Krimphoff, "Perceptual scaling of synthesized musical timbres : Common dimensions, specificities, and latent subject classes," *Psychological Research* **58**, pp. 177–192, 1995.
10. J. Grey, "Multidimensional perceptual scaling of musical timbres," *Journal of the Acoustical Society of America* **61**(5), pp. 1270–1277, 1977.
11. S. Lakatos, "A common perceptual space for harmonic and percussive timbres.," *Perception & Psychophysics* **62**(7), p. 1426, 2000.

12. C. Krumhansl, “An exploratory study of musical emotions and psychophysiology,” *Canadian journal of experimental psychology* **51**(4), pp. 336–353, 1997.
13. K. Scherer and J. Oshinsky, “Cue utilization in emotion attribution from auditory stimuli,” *Motivation and Emotion* **1**(4), pp. 331–346, 1977.
14. P. Juslin, “Perceived emotional expression in synthesized performances of a short melody: Capturing the listener’s judgment policy,” *Musicae Scientiae* **1**(2), pp. 225–256, 1997.
15. P. N. Juslin and J. A. Sloboda, eds., *Music and emotion : theory and research*, Oxford University Press, Oxford ; New York, 2001.
16. A. Friberg, R. Bresin, and J. Sundberg, “Overview of the kth rule system for musical performance,” *Advances in Cognitive Psychology, Special Issue on Music Performance* **2**(2-3), pp. 145–161, 2006.
17. A. Gabrielsson and E. Lindström, *Music and Emotion - Theory and Research*, ch. The Influence of Musical Structure on Emotional Expression. Series in Affective Science, Oxford University Press, New York, 2001.
18. O. Grewe, F. Nagel, R. Kopiez, and E. Altenm
"uller, “Emotions over time: Synchronicity and development of subjective, physiological, and facial affective reactions to music,” *Emotion* **7**(4), pp. 774–788, 2007.
19. E. Schubert, “Modeling perceived emotion with continuous musical features,” *Music Perception* **21**(4), pp. 561–585, 2004.