Audio Engineering Society

Convention Paper 10183

Presented at the AES 157th Convention 2024
October 8–10, New York, NY, USA

Physiological measurement of the arousing effect of bass amplification in music

Nicolas Epain[1], Thomas Mouterde[1], Mérové Wallerich[1], Etienne Corteel[1], and Luc Arnal[2]

[1]*L-Acoustics, Marcoussis, France*
[2]*Institut de l'Audition, Institut Pasteur, INSERM, Paris, France*

Correspondence should be addressed to Nicolas Epain (nicolas.epain@l-acoustics.com)

ABSTRACT

Music's amazing ability to evoke emotions has been the focus of various scientific studies, with researchers testing how different musical structures or interpretations impacted the emotions induced in the listener. However, in the context of amplified music, little is known about the influence of the sound reinforcement system. In this study, we investigate whether the amount of low-frequency amplification produced by a sound system impacts the listener's arousal. We organized two listening experiments whereby we measured the skin conductance of the participants while they were listening to music excerpts with different levels of low-frequency amplification. Our results indicate that an increase in the level of bass is correlated with a small but measurable rise in electrodermal activity, which is correlated with arousal. In addition this effect seems to depend on the nature of the music.

1 Introduction

The relationship between music and emotion has been recognized throughout history. For centuries, composers and music performers have intentionally used elements like tempo, harmony, and instrumentation to evoke feelings. Listening to music is a complex process whereby different perceptual, neurological and psychological mechanisms play a role. Hence, over the past decades research works have explored the relationship between music and emotions from the viewpoints of psychology, neuroscience, and musicology [1].

Aside from investigating the psychological, perceptual and physiological phenomena that govern the occurrence of emotions when listening to music, most of the research has focused on parameters either related to the musical content or to the listener's culture or experience [2]. In comparison, very few studies have looked at factors relating to the way music is presented to the listener. For instance, the sound system configuration (*i.e.,* the type of loudspeakers and amplifiers used, the speaker positions, etc.) could affect the emotions evoked by a given piece of music. Research in this area is useful to the music industry because it will eventually help professionals provide even more memorable music performances. Live music is of particular interest to the authors of this paper, but we believe that this research could benefit to other aspects of the industry, such as music production.

In this paper, we investigate the influence of a basic sound system parameter, namely the amount of amplification of the low frequencies, on the emotional response to music. We set up two listening experiments, one indoor and a second one outdoor, during which the participants were presented with music stimuli while

Fig. 1: Indoor listening test measurement setup. Top left: the USB buzzer button used for participant feedback. Top right: a Shimmer GSR sensor set up on a participant's hand. Bottom: a participant taking the indoor test.

varying the amount of bass amplification. The emotional response of the listeners was monitored using the electrodermal activity (EDA) method.

EDA, which is at the basis of the polygraph (*a.k.a.* the *lie detector*), has been widely used for emotional measurement [3]. It consists in measuring the skin conductance of the test participants with surface electrodes located on their hand, typically. The reason dermal conductance can be used as an indicator of emotions is that it varies with the release of sweat at the skin surface. When a person experiences emotions (arousal in particular), the brain excites the sympathetic nerve system, which in turn triggers eccrine sweat glands located under the surface of the skin. Note that EDA is generally regarded as a good indicator of emotional arousal (*i.e.*, excitement), but not of emotional valence: in other words, it does not inform about the positive or negative character of the emotional response [4]. Nevertheless, previous studies have correlated EDA with the occurrence of chills and the intensity of the emotional response to music [5]. As our objective is to understand how to produce more intense and memorable music listening experiences, EDA appears to be a relevant measurement modality for this study.

The paper is organized as follows. In Section 2, we describe the setup of the indoor listening test. In Section 3 we present the results obtained in this test. Section 4 presents the setup and results of the outdoor test, and Section 5 discusses the significance and limitations of our results.

2 Indoor listening test experimental protocol

In order to measure the impact of low-frequency amplification on the listener's emotional response to music, we organized a first listening test that took place in an indoor environment. The test was held in an acoustically treated auditorium used for multichannel audio mixing. Participants sat in a comfortable armchair, facing an L-Acoustics *Tahiti* 2.1 system, in a setup that resembled that of music listening in a domestic environment (see Fig. 1). The Tahiti system consisted of two *Syva* speakers, two *Syva Sub* subwoofers and one *LA4X* amplified controller.

2.1 Tested conditions

During the test, participants were presented with excerpts of musical tracks, which were filtered to simulate different amounts of low-frequency amplifications. In the following, we refer to the different versions of the stimuli as "LF condition 1–4", where 1 is the condition with the smallest amount of bass and 4 the one with the largest amount of bass. The frequency responses of the filters used to stimulate the stimuli are illustrated in Fig. 2. LF condition 3 corresponded to the condition where no filtering was applied, *i.e.,* the sound system was used in its default preset. Note that the frequency response of the system is not flat in this condition. In comparison, LF conditions 1 and 2 had

Table 1: List of the tracks used in the listening tests.

Label	Artist / Composer	Title	Genre
Sara	Sarah McLachlan	Angel	Pop/Folk
Glad	Hans Zimmer	Am I not merciful? (Gladiator OST)	Epic/Orchestral
Inter	Hans Zimmer	Where we're going (Interstellar OST)	Orchestral/Electronic
Stein	Rammstein	Rammstein	Metal/Industrial
Mess	NAS	The Message	Rap/Hip-hop
Acid	Ayu Acid	Being yourself	Electronic
Echo	Pink Floyd	Echoes	Prog. Rock
Dear	Azymuth	Dear Limmertz	Jazz Funk

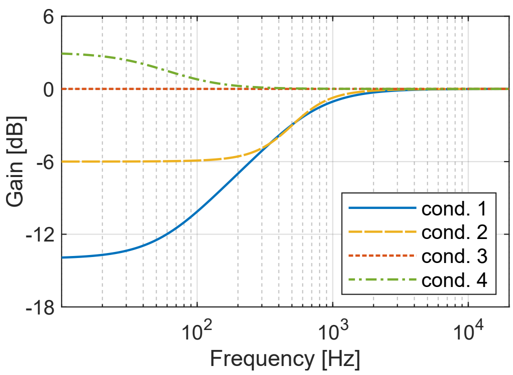

Fig. 2: Frequency equalization applied to the musical tracks to obtain stimuli for LF conditions 1–4.

a reduced amount of low frequencies, with an attenuation of approximately 14 and 6 dB, respectively. On the other hand, with LF condition 4 there was a 3dB boost in the low-end. The different equalization conditions were designed using the *LF Contour* and *Zoom Factor* tools in L-Acoustics' *LA Network Manager*.

2.2 Arousal measurement

Participants were simply asked to sit on the armchair and focus on the music playing on the stereo in front of them. They were handed a buzzer-type joystick (see Fig. 1), which was connected to the test PC via USB and allowed them to provide feedback every time they experienced chills due to the music or simply felt aroused by the music.

Additionally, test participants were equipped with a Shimmer3 GSR+ sensor [6], which was used to record skin conductance for the duration of the experiment. The sensor's electrodes were positioned on the participant's weak hand, as illustrated in Fig. 1. The sensor was connected to the test PC via Bluetooth, which allowed the test operator to monitor the recorded EDA signals in real time and facilitated the synchronization of EDA recordings and stimulus playback.

2.3 Test design and stimuli

Eight musical pieces with diverse genres were used to generate the test stimuli. Within each piece, a time excerpt was selected by the authors so as to provoke a strong emotional response, thanks to the presence of a crescendo, "drop", bouncy rhythm or emotional singing. The duration of the excerpts ranged to 38 to 46 seconds, so that the participants had sufficient time to experience important musical structures such as chord progressions, loops, etc. In the following, we refer to the eight musical excerpts used to generate the test stimuli as the *tracks*. A list of the tracks is given in Table 1.

The eight tracks were filtered using the four filters described in section 2.1, to produce 32 different stimuli. The stimuli were then equalized in loudness using the algorithm described in the ITU loudness measurement standard [7]. However, in the case of track *Glad*, an additional gain of 3 dB was applied to the stimuli because they sounded quieter than the others. This may have been caused by the strong crescendo present in this track. The playback level at the listening position was set at about 90 dBA on average.

The listening test was divided in two measurement sessions, separated by a short pause. Each measurement session started by 1 min of silence, followed by 1 min 30 s of introductory easy listening music, and then a series consisting of 16 of the 32 stimuli. The purpose of the introductory music was for the participants to have a bit of time to relax and to get used to the playback level. The playback level for the introductory music, relative to the stimuli, slowly increased from -12 dB to 0 dB over the course of the 1 min 30 s. The duration of the pause between sessions was at the participant's convenience, but generally lasted a few minutes. Overall, the entire test typically took about 35–40 minutes, pause included.

The order in which the stimuli were played was randomly generated for each participant. However, two constraints restricted the random generation. First, each of the two measurement sessions consisted of exactly two versions (LF condition 1–4) of the eight tracks. Second, a given track could not be played more than once over a series of 5 stimuli. This was done to reduce the participant's weariness resulting from the same tracks repeating several times and maximize their emotional response.

2.4 Test participants and screening

A total of 21 listeners, 13 male and 8 female, participated in the experiment. The age of the participants ranged from 21 to 52 and they all reported having normal hearing abilities. The participants had various backgrounds, however about two thirds practiced music and/or had a profession related to music.

A screening of the data showed that 4 of the participants were EDA "nonresponders", *i.e.,* very little activity could be measured at the surface of their skin. The data measured for these participants was therefore not used is the rest of the study. To identify nonresponders, the skin conductance signals measured for the two measurement sessions were first z-scored. The variance of the signals was then calculated for time windows of duration 10 s, and averaged over every window. The 4 nonresponders presented much lower variances than the rest of the participants.

Note that, according to the literature [3], the proportion of nonresponders in the general population is on the order of 10 to 20%. Therefore, the number of nonresponders detected in this study is not exceptionally large.

3 Indoor listening test results

In this section, we present the results of the indoor listening test.

3.1 Validation of EDA as an arousal measurement modality

Although EDA has been used to measure the emotional response evoked by music in previous studies, it is normally used with very short stimuli because variations in skin conductance can result from different external or internal factors. Listeners could suddenly think of something stressful, such as an important meeting later that day, notice the presence of an object, etc. EDA signals may seem relatively random in that they may feature a number of peaks and variations over the course of a 40 ms music stimulus that do not necessarily occur at the same instants for different listeners or for different listening sessions. Hence, we wanted to make sure that the measured EDA signals presented some form of correlation to the musical content.

We extracted the skin conductance signals corresponding to the different stimuli and, for each track, calculated a skin conductance signature by averaging the signals over every participant and LF condition. Note that, prior to averaging, the skin conductance signals were z-scored because skin conductance values vary greatly from one person to another. Fig. 3 shows the skin conductance signature obtained for track *Stein*, along with the envelope and spectrogram of the corresponding stimulus in LF condition 3. On average, the skin conductance of the listeners increases sharply about 2.5 s after time values t_1 and t_2, which correspond to the song's initial drop and beginning of the vocals, respectively. Therefore, the main arousing musical events in *Stein* are followed by emotional responses, which causes an increase in the listener's skin conductance. The 2.5 s delay that separates the musical events from the increase in skin conductance corresponds to the typical latency reported in the literature [8].

3.2 Listener data processing

Prior to analyzing the test results, the skin conductance signals were processed to obtain one single score per listener and stimulus. Skin conductance signals are often described as the combination of two components: a tonic component and a phasic component. The tonic component is a slowly evolving signal representative of long-term physiological phenomena, such as thermal regulation via sweating. On the other hand, the phasic component consists of faster variations and peaks that are typically related to exterior stimulation, such as auditory or visual events. In order to analyze the impact

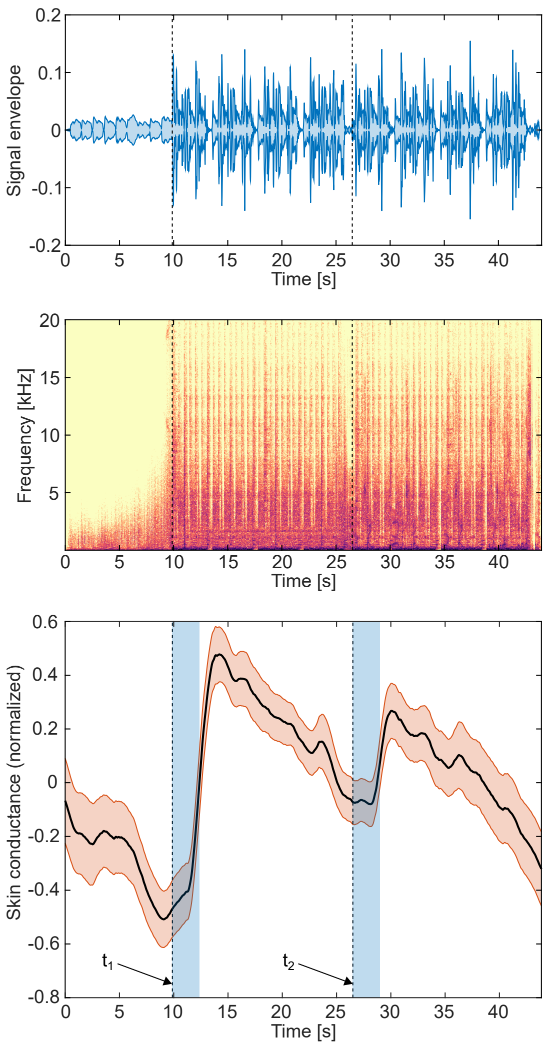

Fig. 3: Relationship between skin conductance variations and musical content. Top: signal envelope for track *Stein* (LF cond. 3). Middle: corresponding spectrogram. Bottom: skin conductance signal measured while listening to track *Stein*, averaged over every test participant and LF condition. The orange area represents the standard error of the mean (standard deviation normalized by the square root of the number of participants). Time values t_1 and t_2 correspond to the "drop" and the moment where the singer starts singing, respectively, and blue areas represent a duration of 2.5 s.

of different factors on the emotional response to music, the phasic component must be extracted from the tonic component.

Various algorithms have been proposed to extract skin conductance peaks and estimate their amplitude, however such methods generally employ some form of thresholding, which poses the risk of leaving out some data and seems more suitable for very short stimuli. Because we needed a score that estimates the electrodermal activity over the entire duration of stimuli, we chose to employ the EDA positive change method (EPC) [9]. This method consists in simply accumulating the positive variations in skin conductance over a period of time, hence summing any occurring peak. A drawback is that a slow (tonic) increase or decrease in conductance can affect the EPC value, but for a 40 s period the impact is relatively small.

After the EPC scores were calculated for each participant and stimulus, a log function was applied in order to increase the normality of the value distribution. Lastly, there were large differences between the EPC values observed for the different participants. However, we were interested in the EPC variations relative to each listener's average electrodermal activity. Therefore, for each participant, the set of EPC values obtained for every stimulus was z-scored. In the following, we refer to the obtained electrodermal activity values as EDA scores.

The data gathered from the feedback button was processed as follows. For each participant and each stimulus, we counted the number of times the feedback button was pressed to obtain one value. Because the number of button presses varied greatly between participants, the sets of obtained values were z-scored for each participant. As well, one participant had to be removed from the dataset because they never pressed the button. In the following, we refer to these scores as FBP (feedback button presses) scores.

3.3 Effect of track repetition

We start our data analysis by examining how the EDA scores evolve over the course of the experiment. This is illustrated in Fig. 4, where the EDA scores are plotted as a function of the stimulus numbers, averaged over every participant and over the two measurement sessions. It appears that the EDA decreases during the measurement, which could be interpreted as the participants being less and less aroused as they listen to

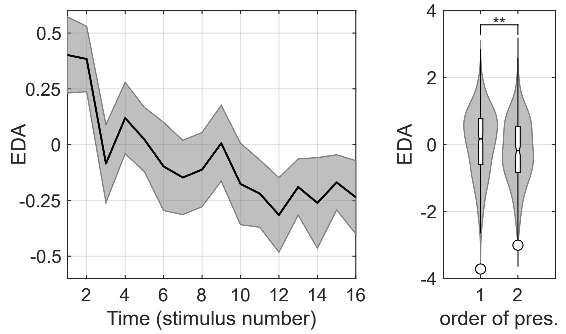

Fig. 4: Indoor test results: influence of stimulus repetition on the measured EDA. Left: EDA during the course of a measurement session, averaged over every participant and session. The gray area represents the standard error of the mean. Right: EDA as a function of tracks order of presentation during a measurement session, averaged over every participant, tracks, and sessions. The EDA measured when a track is heard for the first time is significantly larger than that when it is heard for the second time, as revealed by a t-test ($p < 10^{-2}$).

the sequence of stimuli. As the stimuli were chosen for their relatively high emotion induction power, it is possible that the participants became tired over time.

Another possible reason for this phenomenon is that each track was played twice (with two different conditions) during a measurement session. The emotional response to a given track is likely lessened if this same track has been played a few minutes earlier. To observe this effect, we compared the EDA scores measured the first time a track is played during a session with that measured the second time it is played. A t-test revealed that the value of the EDA score is indeed significantly larger when a track is heard for the first time in a measurement session ($p < 0.01$). In the following we do not consider the effect of time or repetition on the EDA, as it seemed independent of the track or condition.

3.4 Multifactorial analysis

In order to assess the impact of the musical content and LF condition on the participant's emotional response, we ran two-way analyses of variance (ANOVA) with factors "track", "LF condition" and their interaction on both the EDA and FBP scores.

Table 2: Indoor test results: 2-way ANOVA of feedback button presses (FBPs) for factors track, LF condition and their interaction.

Factor	Sum sq.	d.o.f.	Mean sq.	F	p
Track	48.639	7	6.948	4.793	**0.0001**
Cond.	31.299	3	10.433	11.061	**0.0000**
T*C	8.101	21	0.386	0.597	0.9200
Error	454.058	480			

Table 3: Indoor test results: 2-way ANOVA of EDA scores for factors track, LF condition and their interaction.

Factor	Sum sq.	d.o.f.	Mean sq.	F	p
Track	16.256	7	2.322	2.806	**0.0099**
Cond.	10.532	3	3.511	3.917	**0.0140**
T*C	29.101	21	1.386	1.484	0.0799
Error	454.058	512			

The result of the ANOVA on the FBP scores is shown in Table 2. Both the track ($F(7,480) = 4.793$, $p \approx 10^{-4}$) and LF condition ($F(3,480) = 11.061$, $p < 10^{-4}$) have a significant impact on the FBP. However, the effect of their interaction is not significant. In addition, the track has a larger effect size ($\eta^2 = 0.0897$) than that of the LF condition ($\eta^2 = 0.0577$).

The results of the ANOVA on the EDA scores are somewhat similar, as illustrated in Table 3. As is the case for the FBP, both the track ($F(7,512) = 2.806$, $p < 10^{-3}$) and LF condition ($F(3,512) = 3.917$, $p \approx 1.4 \cdot 10^{-2}$) have a significant impact on the EDA. The impact of the track-condition interaction is not significant. As well, the effect size is larger for the track ($\eta^2 = 0.0319$) than for the LF condition ($\eta^2 = 0.0207$).

In summary, there is a measurable impact of the track content and amount of bass amplification on both the feedback provided by the participants regarding their own arousal and their electrodermal activity.

3.4.1 Effect of the musical content

In this section, we examine the effect of the musical content more closely. Fig. 5 shows The FBP and EDA scores as a function of the track. Comparing the data

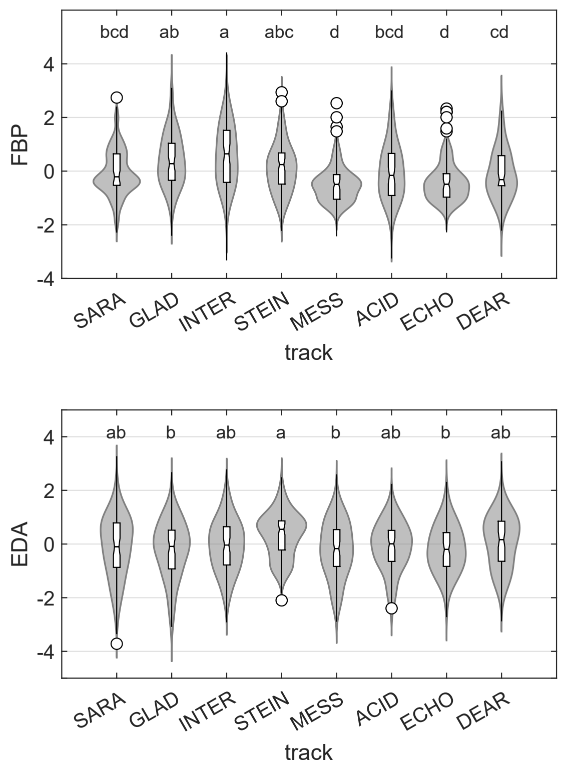

Fig. 5: Indoor test results: influence of the musical track on the participant emotional responses. Top: Number of feedback button presses (FBP) as a function of the track. Bottom: Electrodermal activity (EDA) as a function of the track. Note: the letters in the upper part of the plots indicate statistically significant differences, *i.e.*, groups sharing a letter are not significantly different from each other. Group differences were evaluated using a Tukey-Kramer multiple comparison test (also known as Tukey's HSD [10]).

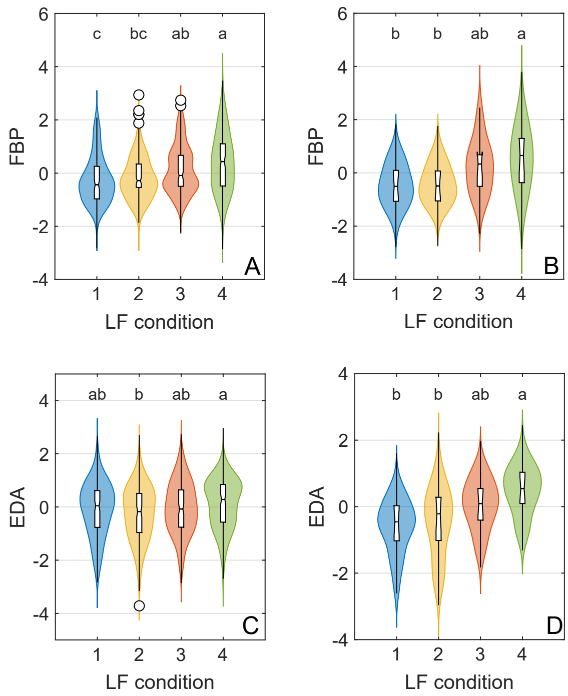

Fig. 6: Indoor test results: influence of bass amplification on the participant emotional responses. A) FBP as a function of the LF condition. B) FBP as a function of the LF condition for track *Acid*. C) EDA as a function of the LF condition. D) EDA as a function of the LF condition for track *Acid*. Other details as in Fig. 5.

obtained for the two measurement modalities, we observe that the FBP values vary more as a function of the track than the EDA value. This was expected given the effect size determined by the ANOVA.

Regarding the impact of the different tracks, based on the feedback of the participants, *Inter* was the most arousing track, while *Mess* and *Echo* were the least arousing. On the other hand, looking at EDA, *Stein* had the biggest impact while *Glad*, *Mess* and *Echo* had the smallest effect. Hence, although the tracks that have the least arousing effect on the listeners seem to be the same in terms of FBP and EDA, the tracks that have the most effect are different. A possible explanation for this result is that the participants may have pressed the feedback button more for tracks that they liked. In other words, the participants' musical preferences could have biased their assessment of how aroused they felt. If this is true, the FBP scores reflect the listener's taste to some extent, instead of being solely related to their arousal.

3.4.2 Effect of the LF condition

We now focus on the effect of the LF condition on the emotional response, which is illustrated in Fig. 6. We first consider the effect on average for every track, presented in Figs. 6A and 6C. As is the case for the effect of the track, the effect of the LF condition on the

FBP is larger than on the EDA. Again, this is supported by the ANOVA's effect size values.

We observe that the FBP increases as the amount of bass increases: the FBP values for condition 4 are significantly larger than that obtained for conditions 1 and 2. Conversely, FBPs obtained for condition 1 are significantly smaller than that obtained for conditions 3 and 4. For the EDA, the effect is not as clear: although EDA values obtained for condition 4 are larger, they are significantly larger than that obtained for condition 2 only.

Examining the data for individual tracks using a 1-way ANOVA for factor LF condition, a significant effect was detected only for track *Acid* (EDA and FBP) and for track *Mess*. The data corresponding to track *Acid* is plotted in Figs. 6B and 6D: for the two measurement modalities, the values obtained with LF condition 4 are significantly larger than that obtained for conditions 1 and 2.

In summary, in our test, low-frequency amplification measurably impacted the listener's arousal both from the perspective of the perception (FBP) and from the perspective of their physiological activity (EDA). In addition, the intensity of this effect varied as a function of the musical content.

4 Outdoor listening test

The listening conditions in the indoor test were somewhat similar to that experienced in a living room. However, ultimately we are more interested in the listening experience found in live music concerts. Therefore, we repeated our experiment outdoors with a setup that resembled that of a music festival.

4.1 Experimental protocol

In the outdoor test, a group of three participants sat side by side, facing a mono sound system consisting of four L-Acoustics' *K2* speakers and three *KS28* subwoofers. The speakers were stacked on the ground about 15 meters from the participants. The stimuli were mono versions of the indoor test stimuli, obtained by summing the left and right channels. The stimuli were loudness equalized and the overall playback level at the listening position was set at around 90 dBA.

Because we had observed a negative impact of stimulus repetitions on the EDA in the indoor test, we increased the number of pauses in the experiment. Instead of two measurement sessions separated by one pause, the outdoor test consisted of four sessions and three pauses. Each session was organized as follows: a 30 seconds of silence, followed by an 1 m and 30 s long musical intro and then 8 test stimuli. In each session, each track was played exactly once. The order in which the different LF versions (four LF conditions) of the tracks were played was random. As well, the order of the tracks within one session was randomized. The duration of the pauses was not imposed, but it lasted two to three minutes on average.

As the indoor test EDA and FBP data showed similar trends, we did not set up feedback buttons for the outdoor test participants. Lastly, another difference with the indoor test is that the Shimmer sensors were not connected via Bluetooth. Instead, the skin conductance signals were recorded in the sensor memory and downloaded to a computer at the end of the listening test. In total, 41 people (21 female and 20 male) took part in the test. Their age ranged from 20 to 80 years old.

Table 4: Outdoor test results: 2-way ANOVA of EDA scores for factors track, LF condition and their interaction.

Factor	Sum sq.	d.o.f.	Mean sq.	F	p
Track	24.427	7	3.489	3.762	**0.0007**
Cond.	4.039	3	1.346	1.643	0.184
T*C	29.997	21	1.428	1.400	0.109
Error	1133.766	1184			

4.2 Results

The measured data was processed in the same manner as that recorded in the indoor test. The screening procedure led to the removal of three listeners out of the 41. We first checked whether we still saw an impact of track repetition on the EDA. A one-way ANOVA indicated that the effect of track repetition was not significant ($F(3,1215) = 2.16$, $p = 0.0913$). Hence, the additional pauses between listening sessions may have solved the listener tiredness problem observed in the indoor test.

We then ran a two-way ANOVA on the EDA for factors track and LF condition and their interaction. The

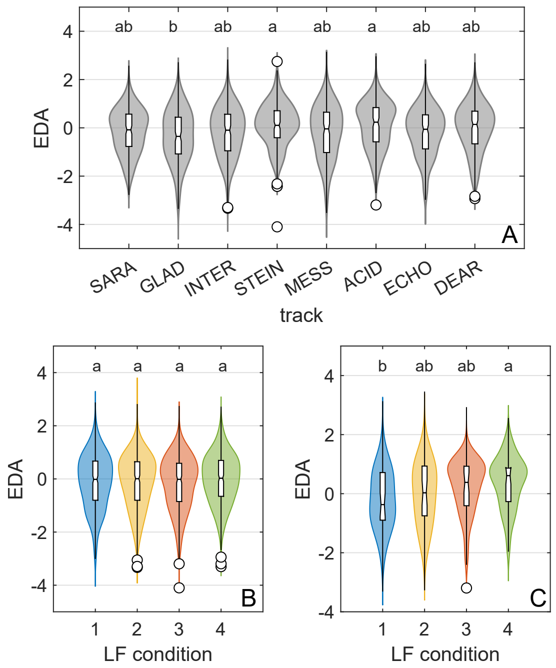

Fig. 7: Outdoor test results: influence of the musical track and LF condition on the participant emotional responses. A) EDA as a function of the track. B) EDA as a function of the LF condition. C) EDA as a function of the LF condition for track *Acid*. Other details as in Fig. 5.

results are presented in Table 4. Contrary to the indoor test, only the track seems to have a significant effect on the EDA ($F(7,1184) = 3.762$, $p = 7 \cdot 10^{-4}$). The corresponding effect size is also smaller than in the indoor test ($\eta^2 = 0.0205$).

The effect of the track on the EDA is plotted in Fig. 7A. The results are somewhat similar to that obtained in the indoor test, but with less contrast between data groups. Like in the indoor test, the track that evoked the strongest reactions in terms of EDA was *Stein*. As well, the track that evoked the weakest reaction was *Glad*, which was also one of the least arousing tracks in the indoor test.

Looking at Fig. 7B, a Tukey-Kramer comparison test confirmed that there was no significant difference between the EDA values obtained for the different LF conditions, on average for every track and participant. However, as shown in Fig. 7C, there is a significant effect for track *Acid*: for this track, the largest EDA are obtained for LF condition 4 and the smallest EDA values for condition 1.

In summary, the results of the outdoor test are similar to those obtained in the indoor test, but with less contrast.

5 Discussion

The results of our study show that the amount of low-frequency amplification has an influence on the arousal experienced by the listener for a given musical track. This effect was detected both from the feedback consciously provided by the listeners (FBP) and observing physiological reactions (EDA).

5.1 Significance of our results

Although the link between low-frequencies and the arousal or energy evoked by music is implicitly attested by certain music production techniques, such as the “bass drop” in EDM, very few studies have properly demonstrated this relationship. A correlation between the presence of very-low frequencies and the amount of dancing has been demonstrated in a recent paper [11], but the frequency range considered was mostly inaudible (below 40 Hz). To our knowledge this is the first time a link between the amount of audible bass and arousal is established using EDA measurements.

Regarding the intensity of the impact of bass amplification on listener arousal, in our data the effect size of the LF condition was only about 30% less than that of the music track, hence it is clearly not negligible. Consequently, care must be taken when designing a sound reinforcement system to make sure enough low-frequency resources are available. As well, this means that the sound mixing approach can impact the audience’s arousal significantly.

Lastly, the influence of the amount of bass on the arousal seems to depend on the musical content. In the test data, the effect was stronger for track *Acid*, which is an EDM track. This is interesting because it is in line with mixing and system design habits observed in the field: certain music genres, including EDM, are regarded as requiring more bass than others (*e.g.*, jazz). However, at present it is not clear why: it could be caused by cultural expectations associated to certain genres, or be related to musical structures, such as the tempo, repetitiveness, etc.

5.2 Limitations of this study

Our results have been obtained with a limited set of music tracks, in conditions that are quite different from that of a live music concert. The outdoor test attempted to approach such conditions, but we set the playback level to about 90 dBA in order to ensure an acceptable sound exposure for the test participants. This made it possible for listeners to be distracted by the frequent airplane flyovers, which partly explains the lower contrast observed in the outdoor test data. Aside from the playback level, it is likely that other factors such as the physical presence of performers or the presence of a dense surrounding audience affect the perceptual and psychological mechanisms that cause the listener emotional response. Hence, in future work, more efforts must be put into establishing experimental protocols allowing to assess the emotional response to music with ecological validity.

Another limitation to our study lies in the comparison between the data obtained via user feedback (FBP) and via EDA measurements in the indoor test. Our data set shows similar trends when it comes to the influence of bass, which indicates that EDA can be used to assess arousal instead of listener report under certain conditions. However, when looking at the influence of the track, the two measurement modalities show different patterns. This could be explained by the fact that listeners are not necessarily conscious of their own emotional arousal, as measured by EDA [12]. In other words, for certain tracks participants may have been physically aroused more than what they thought.

6 Conclusion

In this study, we found evidence of a link between the level of bass amplification in the sound system and the emotional response to music. Namely, bass seems to have an arousing effect on the audience. In future work, we will investigate other parameters related to the sound system design and mixing, such as sound spatialization or dynamic compression.

References

[1] Juslin, P. N. and Västfjäll, D., "Emotional responses to music: the need to consider underlying mechanisms," *The Behavioral and brain sciences*, 31(5), pp. 559–621, 2008.

[2] Garrido, S., "A Systematic Review of the Measurement of Mood and Emotion in Response to Music." *Psychomusicology: Music, Mind, and Brain*, 24, 2015.

[3] Dawson, M. E., Schell, A. M., and Filion, D. L., "The electrodermal system," in J. T. Cacioppo, L. G. Tassinary, and G. G. Berntson, editors, *Handbook of Psychophysiology*, Cambridge University Press, 2000.

[4] Albert, W. and Tullis, T. S., "Measuring Emotion," in W. Albert and T. S. Tullis, editors, *Measuring the User Experience*, Interactive Technologies, pp. 195–216, Morgan Kaufmann, 3rd edition, 2023.

[5] Rickard, N. S., "Intense emotional responses to music: a test of the physiological arousal hypothesis," *Psychology of music*, 32(4), pp. 371—-388, 2004.

[6] "Shimmer3 GSR+ unit," `https://shimmersensing.com/product/shimmer3-gsr-unit/`, nd, accessed: 2024-06-10.

[7] "ITU-R BS.1770-4: Algorithms to measure audio programme loudness and true-peak audio level," 2015.

[8] Posada-Quintero, H. F. and Chon, K. H., "Innovations in Electrodermal Activity Data Collection and Signal Processing: A Systematic Review," *Sensors*, 20(2), 2020.

[9] Leiner, D. J., Fahr, A., and Früh, H., "EDA Positive Change: A Simple Algorithm for Electrodermal Activity to Measure General Audience Arousal During Media Exposure," *Communication Methods and Measures*, 6(4), pp. 237–250, 2012.

[10] Howell, C. H., *Statistical Methods for Psychology, 8th Edition*, Cengage Learning, 2013.

[11] Cameron, D. J., Dotov, D., Flaten, E., Bosnyak, D., Hove, M. J., and Trainor, L. J., "Undetectable very-low frequency sound increases dancing at a live concert," *Current biology*, 32(21), 2022.

[12] Bell, L., Vogt, J., Willemse, C., Routledge, T., Butler, L. T., and Sakaki, M., "Beyond Self-Report: A Review of Physiological and Neuroscientific Methods to Investigate Consumer Behavior," *Frontiers in psychology*, 9(1655), 2018.