

EMuJoy: Software for continuous measurement of perceived emotions in music

FREDERIK NAGEL, REINHARD KOPIEZ, OLIVER GREWE, AND ECKART ALTENMÜLLER
Hanover University of Music and Drama, Hanover, Germany

An adequate study of emotions in music and film should be based on the real-time measurement of self-reported data using a continuous-response method. The recording system discussed in this article reflects two important aspects of such research: First, for a better comparison of results, experimental and technical standards for continuous measurement should be taken into account, and second, the recording system should be open to the inclusion of multimodal stimuli. In light of these two considerations, our article addresses four basic principles of the continuous measurement of emotions: (1) the dimensionality of the emotion space, (2) data acquisition (e.g., the synchronization of media and the self-reported data), (3) interface construction for emotional responses, and (4) the use of multiple stimulus modalities. Researcher-developed software (EMuJoy) is presented as a freeware solution for the continuous measurement of responses to different media, along with empirical data from the self-reports of 38 subjects listening to emotional music and viewing affective pictures.

Since Kate Hevner's early investigations on the perception of emotions while listening to music (Hevner, 1936), there have been many different approaches to the measurement of emotions. For example, Gabrielsson and Lindström Wik (2003) investigated musical expression by describing the verbal reports given by subjects. Most researchers have used distinct adjective scales for the rating of perceived or expressed emotions. Schlosberg (1954) found that such scales can be mapped onto two or three dimensions. Using these methods, self-reported data were collected on distinct and mostly nonequidistant points in time, which were chosen intuitively by the researcher.

The experience of emotion in music and films unfolds over time: This has led to increasing interest in the continuous measurement of perceived emotions, made possible by technological developments since the 1990s. Subjects' responses can now be recorded in real time and synchronized to the stimuli. Schubert (1996, 2001/2002, 2004a, 2004b) was one of the first researchers to develop software that focuses on the perception of the temporal dynamics of emotion.

However, up to now, methods for the recording of continuous responses were based on researcher-developed software solutions, and there was no agreement on the technical means, interfaces, or methods to use. The main aim of this contribution to the ongoing discussion is to propose standardized methods for the continuous measurement of self-reported emotions. The authors have designed new software, EMuJoy, for this purpose. It is freeware and can be distributed and used for research.

Before describing our integrated software solution, we will address four questions: (1) the dimensionality of the

emotion space, (2) technical aspects of data recording, (3) construction of the interface, and (4) the use of multiple stimulus modalities.

Dimensionality of the Emotion Space

Utilizing similarity ratings of affect-descriptive terms, Russell (1980) demonstrated in his circumplex model of affects that such terms can be mapped onto a two-dimensional space. He analyzed the similarity matrix, resulting in a two-dimensional space containing the affective terms. This space was then scaled to fit the dimensions pleasure–displeasure (i.e., valence) and degree of arousal. Some researchers also use a third dimension—namely, dominance (Russell & Mehrabian, 1977). However, arousal and valence appear to be sufficient to explain most of the variance of affective scales (Lang, 1995). Moreover, the use of a computer monitor restricts the software to two dimensions, of which valence and arousal appear to be the most important and universal (Russell, 1983). The use of other dimensions, such as pleasantness and liking, has recently been discussed by Ritossa and Rickard (2004).

With respect to emotions in music, Schubert used Russell's model of two basic emotional dimensions in his *Emotionspace Lab* (2DES; Schubert, 1996). He proved, as Russell did, the validity and reliability of arousal and valence with linguistic expressions, and he also had subjects estimate the expressed emotions in musical pieces. From this investigation, he found that validity and reliability of estimated emotions exist in music as well as in terms or adjectives.

A similar two-dimensional emotion space was also applied to the modality of vision—namely, in the International

R. Kopiez, kopiez@hmt-hannover.de

Affective Picture System (IAPS) by Lang, Bradley, and Cuthbert (1995), which consists of a wide variety of pictures whose impacts on subjects were tested in a large population. Lang et al. used the dimensions valence, arousal, and dominance, just as Russell and Mehrabian (1977) did. Lang et al. displayed these dimensions not in a space, but in the appearance of manikins.

However, a distinction must be made between emotions felt and those perceived. For example, Gabriellson (2001/2002) discussed the relationship between these different types of emotions. He described several possible relationships between the perception and the induction of emotion in listeners, and he concluded that positive as well as negative relationships, or even no relationship, can exist between perceiving and feeling when listening to musical pieces. Further investigation is needed in this important field of emotion research. In our study, we decided to focus on the evaluation of emotions felt while listening to music. One strength of our proposed method is the ability to validate listeners' self-reports about the emotions they felt, which we achieved by measuring psychophysiological parameters, such as skin conductance response.

In addition to the continuous recording of self-reported emotions, recent studies (e.g., Rickard, 2004) have shown that the synchronized recording of peak experiences of music (so-called *chills*) on a separate "chill track" adds important information when recording emotions in real time.

In light of the above-mentioned studies, we chose a two-dimensional emotion space with the dimensions valence and arousal. We proposed to measure both valence and arousal on bipolar scales, with valence ranging from negative to positive and arousal from calming to arousing. We used neither pleasure–displeasure nor sad–happy to elicit valence ratings in order to avoid a bias resulting from preconceived terms.

DATA ACQUISITION

Sample Rate

Data sampling in our paradigm takes place at distinct, equidistant time points. Even if the spectrum of sampled data is band limited, data acquisition can still be considered a process of continuous measurement. Nyquist's sample theorem states that this is possible if the sampling frequency for a band-limited signal is at least twice the frequency sampled. One must therefore leave out frequencies that are higher than half of the sample rate. If there are higher frequencies in the recorded signal, the signal has to be filtered with a low-pass filter. If the filtering were omitted and the signal was reconstructed based on the measurements, this would result in low frequencies not present in the original signal (i.e., *alias frequencies*).

For example, Schubert (2004b) used a sample rate of 1 sample/sec for his continuous measurement, but because the movements of a computer mouse are not necessarily smooth, the signals he measured were not band limited. Therefore, if it is not possible to prefilter the sampled data, the sample rate should be higher for a given data acquisition period. As a sampling rate, we propose 1 sample every 50 msec, which allows for signal frequencies up to 10 Hz.

The recorded data can then be filtered digitally and sampled down; otherwise, they could only be (linearly) interpolated, not simply reconstructed with an identified accuracy.

Using a high sample frequency has numerous advantages. For example, it is important to synchronize the stimulus with the sampled data, and a high sample rate facilitates this process. The high rate also enables the researcher to record self-reported data together with the envelope of the music stimulus, in one file. This technique results in just one time axis and makes it possible to replay the recorded data (e.g., for the ocular inspection of events).

One could argue that the unfolding of emotions over time takes place within a range of 3–5 sec, or even more slowly, so that only a worthless "pseudo-accuracy" would result from a high sample rate. However, the slow realization and unfolding of emotions does not necessarily lead to slow self-reporting; on the contrary, a subject can move from one edge of emotion space to another within a relatively short time. In fact, we have found a high sample rate to be important for a number of reasons. First of all, given the facts above, a single change in affective state will last for a certain amount of time, and the path from one point of emotion space to the other would be lost using a low sample frequency. In particular, spikes in self-reports could not be obtained in such conditions. Second, our software is appropriate for investigating emotional self-reports for stimuli such as advertisements or music videos, in which scenes are displayed for a very short time in conditions that should enable subjects to report their emotions. Third, the software can display or replay stimuli of very short durations. For instance, images can be displayed for fractions of a second, and the on- and offset of the stimulus presentation should then be logged together with the self-report. For this purpose, the necessity of a high sample rate is unquestionable.

Interface Construction for Emotional Responses

In this section, we describe the construction of our user interface for continuous measurement of self-reported data (see Figure 1), which is a further development of the 2DES interface. As in the original 2DES design, we use a pointer on the screen to depict the position within the emotion space. The pointer can be controlled with a joy-

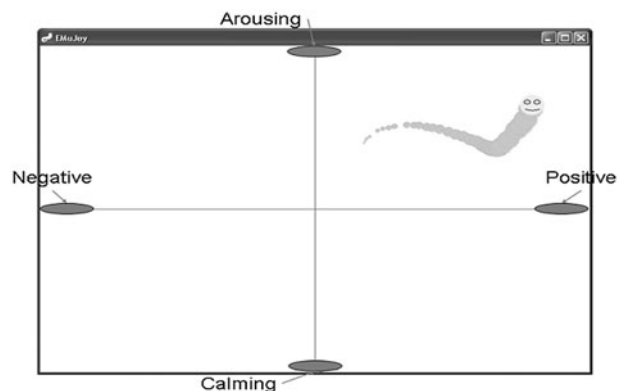


Figure 1. Screen shot of the EMuJoy software interface.

stick or a computer mouse. With our software, in contrast to 2DES, the whole screen can be used for the emotion space; there is no need for the two dimensions to have equal-scaled axes. A change in color of the pointer displays reported chills after a mouse click or a response via any other human–computer interface. The pointer is in the shape of a worm with a tail and a face, which makes the handling very intuitive and uncomplicated: By looking at the face, the subject gets intuitive feedback within the emotion space on his or her self-report. To illustrate the dynamics of the responses, the tail traverses the trajectory of the previous points in the emotion space. This design is related to that of the IAPS, which also uses facial icons for its “Self-Assessment Manikin” (Lang et al., 1995). Schubert (2004a) used generated faces to support self-reporting of emotions in EmotionFace, an alternative to his 2DES.

The facial icon consists of two eyes and a mouth. The eyes open and close according to the vertical “arousal” dimension (defined in Equation 1). The mouth illustrates the horizontal “valence” axis by raising or lowering at the corners (see Equation 2 at the bottom of this page).

$$x^2 + \frac{y^2}{\left(\frac{a-1}{2} + 1\right)^2} = 1; \quad -1 < a(\text{rousal}) < 1 \quad (1)$$

The values for arousal and valence used in both of the formulas are normed to range from -1 to 1 . The coordinates x and y are relative to the positions of the mouth and eyes in the pointer on the screen.

The tail not only shows the previous positions of the pointer, it gives the impression of dynamic movement. It tapers out over the preceding points in circles of decreasing diameter, creating the impression over time of a “moving worm” (see Figure 1).

Training and the Use of Multiple Modalities

When subjects are required to express emotions within an emotion space, especially with computer software, they must first be trained in the handling of the system, in order to become comfortable with it. During a training session, subjects become familiar with the particular demands of a dual-task situation (simultaneous listening or viewing and creation of a self-report using interface movements within the two-dimensional emotion space). To familiarize subjects with the emotion space yet avoid priming effects, we propose using different stimuli in training than during the experiment, preferably stimuli in another modality.

For example, for training before our investigation of perceived emotions in music, we used pictures from the IAPS. We selected pictures that provided neutral and ex-

treme examples of arousal and valence (high/high, high/low, etc.). The pictures were displayed in the background of the emotion space, and the subjects learned to use the software by expressing their perceived emotions while looking at the pictures. The measured values could then be compared with those published by Lang et al. (1995). Video files could also be played in the background of the emotion space. Therefore, subjects would not need to pay attention to different monitors or interfaces; both the presentation of stimuli and the self-reports can be realized within one device.

THE EMuJoy SOFTWARE SOLUTION

EMuJoy stands for “Emotion measurement with Music by using a Joystick.” It enhances Schubert’s 2DES software in three ways: First, it enables subjects to give self-reports for different media in real time with a computer mouse, joystick, or any other human–computer interface. The ability to use a joystick can be an advantage, because joysticks have a return spring to automatically realign them in the middle of the space. This contrasts with Schubert’s software, with which only a mouse can be used.

Second, EMuJoy has numerous useful features for the researcher. It can be completely controlled over the Internet, for instance, so any audiovisual stimuli can be presented to subjects without any personal contact with researcher. Even data acquisition can take place over the Internet.

Third, EMuJoy is a platform-independent solution that is also open to modification, such as customization for further multimodal stimuli.

Handling and Remote Control

The stimuli for a particular subject do not have to be selected before the experiment starts, so they can be chosen, for example, to adapt to the subject’s changing moods. The software provides network access so that stimuli can be selected and started remotely over the Internet. The architecture of the network communication is illustrated in Figure 2.

The main application runs on a subject’s computer, where the stimuli are presented. The researcher controls the application via remote control software that enables selection, starting and stopping, and so forth to be performed online. The communication between the server and the local application is based on the TCP protocol. The researcher can choose any media that are available in the network (in the local area network as well as over the Internet). The separate parts of the program are usually started on different computers; however, it is also possible for all components to run on only one computer.

$$x \in \{0, 1, 2, 3, 4\}; \quad y = \begin{cases} 3(2v-1), & x = 0 \wedge x = 4; \\ 2v-1, & x = 1 \wedge x = 3; \\ 0, & \text{other;} \end{cases} \quad -1 < v(\text{alence}) < 1 \quad (2)$$

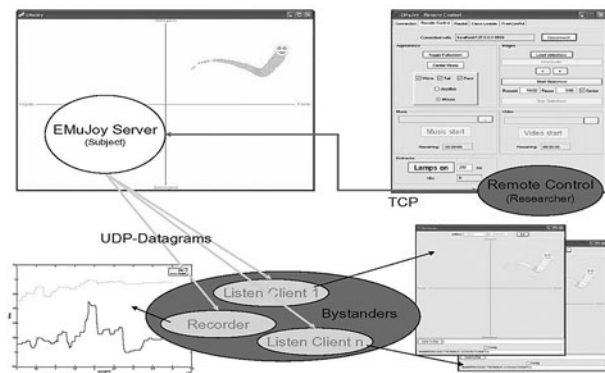


Figure 2. Design of EMuJoy's network communication. All parts of the software and their interconnections are displayed.

Communication then takes place via “localhost” (i.e., by loopback).

A second aspect of the network communication is the use of “bystander” clients, which can run on several hosts. They receive the data from the subjects’ actions, which can then be recorded onto hard disc or simply viewed, for instance for educational purposes. These clients receive the data via the UDP protocol, which has the advantage that packets can be transmitted in real time. This technique is also used for Internet broadcasting.

Analog Output

As we mentioned (and like the 2DES), the EMuJoy software can record the acquired data to disc. The great advantage of EMuJoy, however, is that it allows analog output, which requires a digital-to-analog converter card. EMuJoy contains a simple application programming interface with which a library can easily be programmed to allow any card to provide the appropriate output. The analog output can also be transmitted via the parallel port, limiting the resolution of the output. We have already implemented a driver for the Data Translation 2821 card (Data Translation, www.datx.com). With this output, there is no difficulty synchronizing stimuli with the self-reported data. Thus, all information can be recorded together within one system.

The output includes values for both axes and an additional trigger track. This track contains, for example, data about experienced chills, start and stop of media, and so forth.

Multimedia Stimuli and the Appearance of the Software

A great advantage of the proposed interface is its openness to the use of multimodal stimuli. In addition to musical stimuli, images or films can also be displayed in the background of the emotion space, so that visual as well as acoustical stimuli can be investigated. This makes EMuJoy particularly suited to cross-modality research. Numerous options enable the adaptation of the interface surface to meet personal needs. The display order of pictures can be prepared with text files, and groups of pictures can be presented with or without pauses for set amounts of time. For instance, oddball paradigms can easily be used. Even software modifications affecting the display time or changing the modality of the presented stimuli (e.g., first an auditory, then a visual stimulus, then both together) can be made.

The user has the option of modifying the appearance of the space. For example, all colors can be personalized, the labels can be changed, and the face and the tail can be switched off. The software is released under the GNU general public license (GNU, 1991) and can therefore be modified in other ways; the source is open.

SAMPLE APPLICATION: AFFECTIVE PICTURES AND EMOTION-INDUCING MUSIC

In an exploratory experiment that contributed to a research project concerning strong emotions while listening to music, subjects listened to musical pieces or viewed pictures from the IAPS. We asked them to report the emotions they felt. This experiment differed from several other previous studies, in which subjects were asked to report the emotions they recognized in the stimuli.

Materials and Procedure

The IAPS consists of about 800 affective pictures (Lang et al., 1995). Ten of these pictures were selected to cover the range of the emotion space. In addition, seven musical pieces were selected for another study. Four of the pieces were also chosen to define the edges of the emotion space. The content and the numbers of the pictures we chose are displayed in Table 1.

We present the results of the emotional self-reports while subjects listened to the following two musical pieces: “Making Love Out of Nothing at All” by Air Supply (1983) is an example of pop music. This piece of music was previ-

Table 1
Content of Pictures From the IAPS, With Their Expected Rough Mean Ratings Within the Emotion Space

	Low Valence	Neutral	High Valence
High arousal	cut through a throat (3071) scene of violence (3530)		rafting-sport (8370) erotic female (4240)
Neutral		spoon (7004) basket (7010)	
Low arousal	graveyard (9220) scene in a hospital (2210)		tiny rabbit (1610) child (9070)

Note—IAPS numbers are indicated in parentheses.

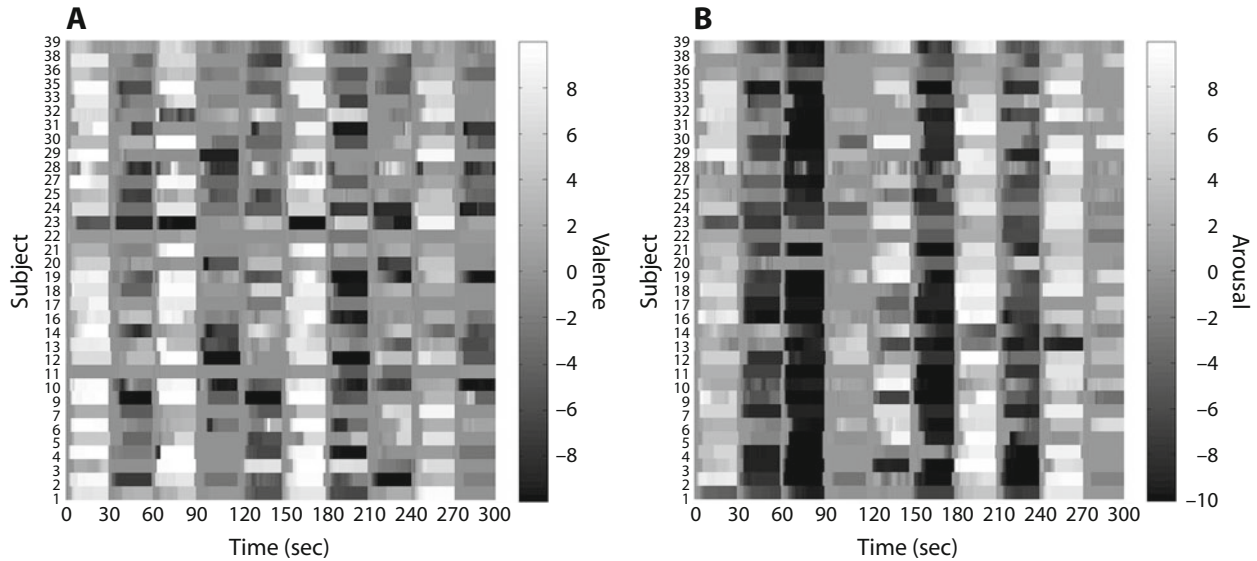


Figure 3. Self-reported valence (A) and arousal (B) of subjects while viewing pictures from the IAPS. Data are depicted from a bird’s eye view. Each row denotes an individual (numbered from 1 to 39.) On the (horizontal) time axis, the unfolding of emotions over time is visualized. The IAPS pictures changed every 30 sec, leading to similar reactions interindividually (shown in the columnar pattern of the graphs). The degrees of valence and arousal are encoded in grayscale, from pure black for -10 to pure white for $+10$, as shown in the color bars.

ously used in a study by Panksepp (1995) and was shown to arouse strong emotional reactions. “A Skull Full of Maggots” by Cannibal Corpse (1990) is a death metal song; it was rated high on the dimension of arousal and low on the valence dimension.

While the music was played or the pictures were displayed, subjects reported in real time the emotions they felt. These self-reports were recorded and compared among all the subjects. A total of 38 subjects (mean age = 38 years, $SD = 16$, range = 11–72; 29 females and 9 males, 33 right- and 5 left-handed) participated in the experiment.

Results and Discussion

Rating the IAPS samples. Figure 3 shows the self-reports of valence and arousal for the affective pictures over time, displayed from a bird’s-eye view. “Columns” of 30 sec appear (i.e., similar coloring in a vertical direction). From these results, it seems that the self-reports were quite similar among all of the subjects.

Table 2 shows the relationship between Lang et al.’s (1995) data and ours. The median rating for each picture is shown, calculated over subjects from the medians of the self-reports over 30 sec. It is remarkable that the data are

Table 2
Results of Self-Reports While Viewing Pictures From the IAPS

No.	IAPS No.	Val (IAPS)	Aro (IAPS)	Val (IAPS, Adapted)	Aro (IAPS, Adapted)	Val (EMuJoy)	Aro (EMuJoy)
1	8370	7.77	6.73	6.93	4.33	3.56	6.39
2	9220	2.06	4.00	-7.35	-2.50	-5.02	-1.13
3	3071	1.88	6.86	-7.80	4.65	-9.15	5.54
4	7010	4.94	1.76	-0.15	-8.10	0.03	-0.90
5	9070	5.01	3.63	0.02	-3.43	0.82	-1.35
6	3530	1.80	6.82	-8.00	4.55	-8.04	6.86
7	1610	7.69	3.98	6.73	-2.55	6.55	-2.55
8	2205	1.95	4.53	-7.63	-1.18	-5.06	0.00
9	4220	6.60	5.18	4.00	0.45	5.64	2.51
10	7004	5.04	2.00	0.10	-7.50	0.06	-0.17

Note—The first column contains the position in the presentation order, followed by the IAPS number and the IAPS valence (Val) and arousal (Aro) ratings. The final two columns contain the raw data collected with the EMuJoy software, presented as the medians over all subjects of the median ratings for each picture over 30 sec. The two columns before them present the adapted IAPS values when those ratings are projected from Lang et al.’s (1995) emotion space into ours. This conversion takes place as described in Equations 3–6.

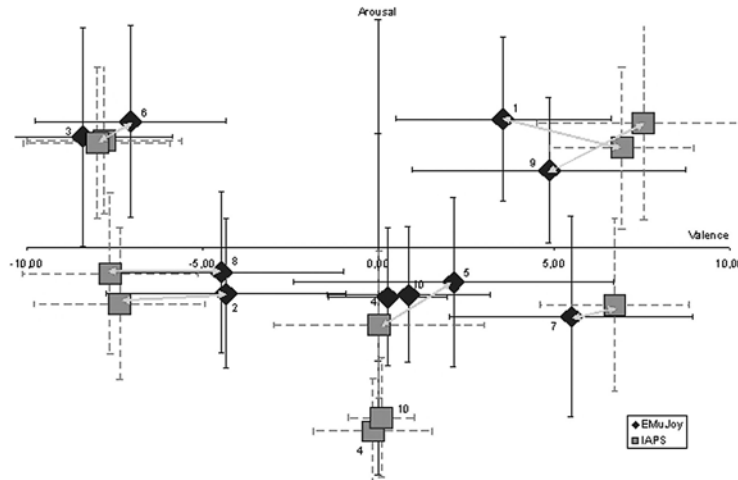


Figure 4. Comparison of the IAPS data and those acquired using the EMuJoy software. Means and standard deviations are displayed, with corresponding data for the same pictures connected by arrows. Note that the means for Pictures 4 and 10 are not very close to each other. These were the neutral images (a basket and a spoon), and the difference in ratings may be due to the different scaling of arousal in the two spaces.

so similar, despite the different instructions and different underlying scales of our two studies. That is, Lang asked his subjects to report all at once about the emotion a picture expressed, but we asked subjects to report continuously about the emotions they felt. Lang et al.’s original arousal data are all slightly higher than ours. This can be explained by the fact that they used a unipolar scale ranging from 1 to 9, whereas we used a bipolar scale ranging from -10 to 10 . In order to allow comparisons, Lang et al.’s ordinal scaled data were projected onto our interval scale—that is, the valence and arousal ratings were projected from their 1 to 9 scale to our scale of -10 to 10 . For that purpose, empirical means and standard deviations were calculated from the data collected with the EMuJoy software.

For the purpose of comparing our data to those collected by Lang et al. (1995), the medians of our data, over 30 sec for each picture, had to be compared with data that were sampled at distinct time points (one data point per picture). Since Lang et al. based their method on the assumption of normally distributed data, we followed this method by displaying data parametrically; for example, we display means with standard deviations. However, we assumed that, because of our limited sample size, the test for normal distribution in our experiment would not be significant. However, if Lang et al. are right, our data would converge toward a normal distribution when used with larger populations. This means that at the current state of research, for reasons of comparability, we assume that our data will have the same distribution as Lang’s.

Equations 3–6 show the formulas for the projection and conversion of the empirical standard deviations. Means as well as standard deviations (*SDs*) have to be translated and scaled. In the equations below, the new mean is denoted by χ , the *SD* by ζ , and empirical values are overlined.

Projection: χ

$$\chi := \frac{x-1}{8} 20 - 10 = \frac{5x-25}{2} \quad (3)$$

Empirical Mean: $\bar{\chi}$

$$\bar{\chi} := \frac{5\bar{x}-25}{2} \quad (4)$$

Empirical standard deviation for x : \bar{s}

$$\bar{s} := \sqrt{\frac{1}{n-1} \sum_{i=1}^n (\bar{x} - x_i)^2} \quad (5)$$

Empirical standard deviation for χ : $\bar{\zeta}$

$$\begin{aligned} \bar{\zeta} &:= \sqrt{\frac{1}{n-1} \sum_{i=1}^n (\bar{\chi} - \chi_i)^2} \\ &= \sqrt{\frac{1}{n-1} \sum_{i=1}^n \left(\frac{5\bar{x}-25}{2} - \frac{5x_i-25}{2} \right)^2} \\ &= \sqrt{\frac{5}{2} \bar{s}} \end{aligned} \quad (6)$$

The comparisons of the valence and arousal data from the two studies within our emotion space, ranging from -10 to 10 , are displayed in Figure 4. Note the close similarity of locations for the self-reports in our space and in Lang et al.’s (1995) adapted space.

Rating musical pieces. In the auditory domain, the results for all subjects were not as homogenous as in the visual domain. Figure 5 shows the results of self-reporting as our subjects listened to the two musical pieces. Overall valence and arousal were similar for all subjects with both

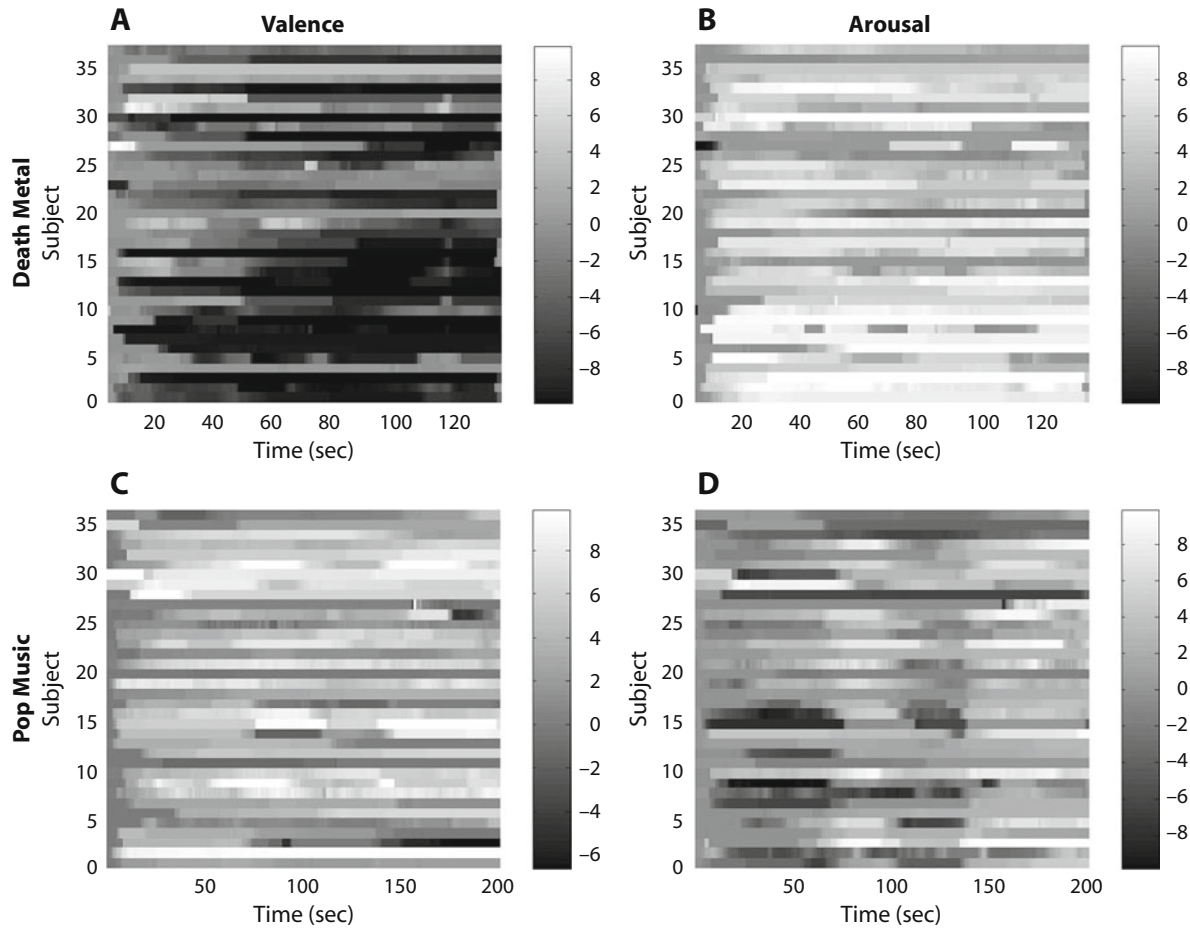


Figure 5. Self-report data for subjects listening to two musical pieces, displayed from a bird's eye view (as in Figure 3). The two upper panels display emotional self-reports while listening to the death metal song “A Skull Full of Maggots” by Cannibal Corpse. In contrast to Figure 3 (and to panel D at the lower right), no columns can be identified in these data, whereas the horizontal patterns demonstrate high individual constancy. Overall tendencies toward aversion (A) and high arousal (B) are ascertainable among all subjects. The lower panels display self-reports while listening to the pop song “Making Love Out of Nothing at All” by Air Supply. As can be seen in panel D, changes in self-reported arousal that are consistent between individuals took place around 70, 100, and 140 sec.

musical pieces. In contrast to the results with the pictures, however, “columns” of similar ratings only become visible in Figure 5D. In the other three panels, these blocks cannot be separated. This means that between subjects, the absolute value of self-reporting in small time windows is more heterogeneous with music than with strong affective pictures. In Figure 5D, the self-reported data do show blocks of similar reactions in the arousal dimension; column boundaries can be identified around 70, 100, and 140 sec. These borders are related to changes in loudness; more precisely, elevated loudness is related to high arousal. In the death metal piece, no similarities in arousal can be found, and there is no relation to loudness or any other obvious features of the piece. At the current state of investigation, we assume that it is possible to find musical pieces with self-reports that are highly consistent intersubjectively, as well as others with highly subject-dependent reports. However, this study was limited to 38 subjects. Collecting data from much larger populations—for example, via an

Internet survey—might demonstrate similarities within all kinds of musical styles, and similar responses to distinct pieces might appear within groups with the same musical interests, age, and socialization. We presume that music may also exist that induces general emotional reactions, as has been observed in other modalities.

An additional explanation of the heterogeneity in the auditory domain, in contrast with the findings by Lang et al. (1995), can be found in Schubert and Dunsmuir (1999). Those authors allude to effects such as an “after-glow” and different time lags in the emotional responses toward musical stimuli. An important factor might be low sensitivity for the relatively “weak” stimuli provided by musical pieces, in comparison with the extremely emotional (violent or salacious) pictures from the IAPS.

In accord with Schubert (1999), we found continuous self-reports within a two-dimensional emotion space to be appropriate for measuring emotions. Even emotions felt, in contrast to emotions perceived, can be depicted via

this technique, particularly when subjects view emotion-inducing pictures. The same emotion space is also appropriate for music. Investigating the perception of music is therefore reliable and valid, as Schubert (1999) showed.

The heterogeneity in the perception of music might be explained by the fact that by using long musical excerpts, distinct emotions change frequently over time. In the visual domain, on the other hand, the same pictures were displayed for long periods of time (30 sec). In the acoustical domain, it is conceivable that weak emotions are “overwritten” by new ones before they become conscious. Thus, only emotions whose intensity is above a certain threshold would be reported, as shown in the block transitions in Figure 5D. Finding such thresholds could be the subject of further investigations.

CONCLUSION AND OUTLOOK

With the EMuJoy software, we have further developed the measurement of continuous self-reporting, thus proving the need for a standard of investigating emotional self-reports with emotional multimodal stimuli. Because of the open source nature of the software, there are no limits in how it could be further developed and adapted for any needs.

Hopefully, other researchers will adapt the software to their needs and contribute new features. Our software is also not restricted to nonprofit uses, and new applications are particularly likely within advertisement or the music industry.

AUTHOR NOTE

This work was supported by the Deutsche Forschungsgemeinschaft (Grant Al 269-6) and the Center for Systemic Neurosciences, Hanover, Germany. The EMuJoy software is freeware written in Java, can be used on any platform, and works as well on PCs as on UNIX workstations. It can be downloaded from musicweb.hmt-hannover.de/emujoy/. Correspondence relating to this article may be sent to R. Kopiez, Hanover University of Music and Drama, Emmichplatz 1, 30175 Hanover, Germany (e-mail: kopiez@hmt-hannover.de).

REFERENCES

- AIR SUPPLY (1983). Making love out of nothing at all. *On The very best of Air Supply* [CD]. San Francisco: Musicrama.
- CANNIBAL CORPSE (1990). A skull full of maggots. *On Eaten back to life* [CD]. Simi Valley, CA: Metal Blade Records.
- GABRIELSSON, A. (2001/2002). Perceived emotion and felt emotion: Same or different? *Musicae Scientiae*, *6*(Special Issue 2001/2002), 123-147.
- GABRIELSSON, A., & LINDSTRÖM WIK, S. (2003). Strong experiences related to music: A descriptive system. *Musicae Scientiae*, *7*, 157-217.
- GNU (1991). GNU general public license. Available at www.gnu.org/licenses/gpl.txt.
- HEVNER, K. (1936). Experimental studies of the elements of expression in music. *American Journal of Psychology*, *48*, 246-268.
- LANG, P. J. (1995). The emotion probe: Studies of motivation and attention. *American Psychologist*, *50*, 372-385.
- LANG, P. J., BRADLEY, M. M., & CUTHBERT, B. N. (1995). *International affective picture system (IAPS): Technical manual and affective ratings*. Gainesville: University of Florida, Center for Research in Psychophysiology.
- PANKSEPP, J. (1995). The emotional sources of “chills” induced by music. *Music Perception*, *13*, 171-207.
- RICKARD, N. S. (2004). Intense emotional responses to music: A test of the physiological arousal hypothesis. *Psychology of Music*, *32*, 371-388.
- RITOSSA, D. A., & RICKARD, N. S. (2004). The relative utility of “pleasantness and liking” dimensions in predicting the emotions expressed by music. *Psychology of Music*, *32*, 5-22.
- RUSSELL, J. A. (1980). A circumplex model of affect. *Journal of Personality & Social Psychology*, *39*, 1161-1178.
- RUSSELL, J. A. (1983). Pancultural aspects of the human conceptual organization of emotions. *Journal of Personality & Social Psychology*, *45*, 1281-1288.
- RUSSELL, J. A., & MEHRABIAN, A. (1977). Evidence for a three-factor theory of emotions. *Journal of Research in Personality*, *11*, 273-294.
- SCHLOSBERG, H. (1954). Three dimensions of emotion. *Psychological Review*, *61*, 81-88.
- SCHUBERT, E. (1996, August). *Measuring temporal emotional response to music using the two-dimensional emotion space*. Paper presented at the 4th International Conference on Music Perception and Cognition, Montreal, Canada.
- SCHUBERT, E. (1999). Measuring emotion continuously: Validity and reliability of the two-dimensional emotion-space. *Australian Journal of Psychology*, *51*, 154-165.
- SCHUBERT, E. (2001/2002). Correlation analysis of continuous emotional response to music: Correcting for the effects of serial correlation. *Musicae Scientiae*, *6*(Special Issue 2001/2002), 149-171.
- SCHUBERT, E. (2004a, July). *EmotionFace: Prototype facial expression display of emotion in music*. Paper presented at ICAD 2004, the 10th Meeting of the International Conference on Auditory Display, Sydney, Australia.
- SCHUBERT, E. (2004b). Modeling perceived emotion with continuous musical features. *Music Perception*, *21*, 561-585.
- SCHUBERT, E., & DUNSMUIR, W. (1999). Regression modelling continuous data in music psychology. In S. W. Yi (Ed.), *Music, mind, and science* (pp. 298-352). Seoul: Seoul National University Press.

(Manuscript received September 20, 2005;
revision accepted for publication January 28, 2006.)