Beyond Demand: Investigating Spontaneous Evaluation of Chord Progressions with the Affective Priming Paradigm

Mira Müller
University of Leipzig, Leipzig, Germany

Julian Klein
University of Arts Berlin, Berlin, Germany

Thomas Jacobsen
Helmut Schmidt University Hamburg, Hamburg, Germany

We assume that evaluative processes in response to musical stimuli can occur spontaneously without explicit demand, and that these responses are important for the emergence of emotions evoked by music. Two versions of the affective priming paradigm served to study spontaneous evaluation of music. In Experiment 1, a lexical decision task (LDT) and in Experiments 2 and 3, an evaluative decision task (EDT) was employed. A total of 20 original four-part, five-chord piano sequences with no specified harmonic resolution were used as primes. During the LDT, congruency in valence of prime-target pairs did not affect response times to the targets. However, for the EDT, significant effects of priming were obtained, indicating that spontaneous evaluations of primes must have occurred. No moderating influences of music expertise or any other person variable on spontaneous evaluation were observed. The diverging results of LDT and EDT point to the possibility that spontaneous evaluative processes are sensitive to context manipulations. Results are discussed with reference to harmonic and semantic priming studies.

Received May 15, 2009, accepted April 14, 2011.

Key words: spontaneous evaluation, affective priming, music aesthetics, expertise, music induced emotions

Aesthetic evaluative responses to music are an integral part of music reception. Thus, a growing body of research exists approaching this phenomenon from various angles. Foremost, researchers have focused on measuring explicit evaluative judgments formed by listeners when instructed to attend to and evaluate musical stimulus material (e.g., Altenmüller, Schürmann, Lim, & Parlitz, 2002; Müller, Höfel, Brattico, & Jacobsen, 2010; Orr & Ohlsson, 2005; Peretz, Gaudreau, & Bonnel, 1998). Beyond situations like this, which imply that evaluations are overt and prompted, a multitude of other listening situations that lead listeners or performers to evaluate music are conceivable. Supposedly, in many of these situations, music is not the primary stimulus monopolizing all available attentional resources. In most people's daily lives, music often plays in the background, accompanying other primary tasks[1] (Sloboda & O'Neill, 2001). Situations in which music acts as a secondary stimulus have been found to be more common than situations in which music takes the role of primary stimulus, being attended to by undistracted listeners[1] (Sloboda & O'Neill, 2001). Hence, an imbalance exists: listening situations set up in most music processing studies are not the music listening situations that occur most frequently in people's daily lives. To counter this trend, the current study recreates one of the most common listening situations, by measuring spontaneous evaluative responses in a situation in which music is presented as a secondary stimulus.

Lately, the fundamental role of aesthetic evaluations of music has been emphasized in theoretical frameworks that focus on emotions felt in response to music (Scherer & Zentner, 2008; Zentner, Grandjean, & Scherer, 2008). In these frameworks, affective reactions—for example, feeling awed, dazzled, moved, and thrilled—have been clearly linked to evaluative processes. This means evaluative responses are an essential prerequisite for emotional responses to occur. It is a long existent idea that emotions in general are fuelled by spontaneous evaluative processes that quickly assign positive or negative valence to objects without being explicitly prompted

[1] Participants of Experiment 3 of the present study on average declared that they actively listened to music for less than an hour per day ($M = 0.89$, $SD = 0.66$). In contrast, they passively listened to music for 2.83 hours per day ($SD = 1.77$). This difference was significant in a paired-t-test, $t(38) = -6.80$, $p < .001$.

Music Perception volume 29, issue 1, pp. 93–108. issn 0730-7829, electronic issn 1533-8312. DOI: 10.1525/mp.2011.29.1.93

(Ortony, Clore, & Collins, 1988). Thus, it is interesting to study the evaluative processes that occur spontaneously in response to music, in order to contribute to the discussion of mechanisms responsible for the origination of emotions during music listening.

Fast and spontaneously occurring evaluations have been intensively studied, with various response-time paradigms; for example, the affective priming paradigm (Bargh & Chartrand, 1999; Fazio, 2001; Murphy & Zajonc, 1993; Wittenbrink, 2007). With this procedure, an index is obtained of the inclination to respond positively or negatively to a given stimulus, in the absence of explicit demand to evaluate the stimulus. More precisely, response latencies in categorization tasks performed under time pressure form the basis for inferences about spontaneously occurring evaluative processes.

In the field of music research, Sollberger, Reber, and Eckstein (2003) have studied spontaneous evaluation at the single chord level relying on the affective priming paradigm. They used two consonant chords as positive and two dissonant chords as negative prime stimuli, and positive and negative words as target stimuli. Participants were presented with prime-target pairs either matching or not matching in valence, and they were instructed to categorize the words according to their valence. As a result, Sollberger et al. (2003) observed significantly faster response times to matching than to mismatching trials, which led to the conclusion that the chords must have undergone a process of spontaneous evaluation as the outcome of this process affected responses to the targets.

This study operated at the level of single chords. Music, however, usually consists of more than one sound event. These sound events are organized in time and are not heard and evaluated as standalone entities, independent of each other. Valence in music, in fact, often results from how these events are interconnected. For example, five chords that are assigned positive or neutral polarity if played in isolation might be evaluated negatively if they are played in close succession. The creation of strong polarities in this manner—with the constituting elements presented one after the other instead of exactly at the same time—is a characteristic unique to music.

The claim that this mechanism is the source for musically induced emotion goes back to Meyer (1956). It has been empirically addressed by Steinbeis, Koelsch, and Sloboda (2006), who replaced individual chords in selected music passages with chords considered to not fit as well in the context in which they appeared. Subsequently, they measured behavioral and physiological responses to the original and altered material. As one behavioral measure, continuous subjective ratings of emotionality were obtained from the participants while they listened to the passages. However, no local changes on this measure of affective response were observed for the altered chords even though the general reactions to the entire passages; that is, the final ratings of overall emotionality at the end of each piece, were affected by the alterations (Steinbeis et al., 2006). The sequences with the most unexpected harmonic events received the highest emotionality ratings at the end of each piece. The ERP data reported by the same authors hints at evaluative processes taking place locally, but no definite conclusions were drawn. Thus, the issue of whether harmonic relations are affectively evaluated on a local level warrants further empirical attention.

The Present Study

As pointed out earlier, many common listening situations lack explicit prompts to evaluate the music that is heard, but nevertheless give rise to emotions that have been termed "aesthetic" in previous reports (Scherer, 2004; Scherer & Zentner, 2008). Their emergence implies that evaluative processes have previously occurred. Thus, it is conceivable that spontaneously occurring evaluative responses underlie these emotions. In the scope of the present study, we examine these spontaneous processes and discuss our results in reference to musically induced emotions.

Extending the findings of Sollberger et al. (2003) and Steinbeis et al. (2006), we tested the assumption that evaluative processes occur spontaneously in response to musical stimuli—even in the absence of explicit demand—by employing the affective priming paradigm as a covert method. Three experiments in total were conducted using two different versions of this paradigm. As musical prime stimuli, chord progressions were used in the experimental conditions and single chords in the control conditions.

In Experiment 1, a nonaffective task was introduced, diverting attention from the valence dimension present in the target stimuli. As target stimuli, German words and pseudowords were presented which had to be categorized according to their lexical status (Lexical Decision Task (LDT)). Target stimuli that were not pseudowords, but fell into the category of words in the German lexicon, either had a positive or negative valence. In contrast to Experiment 1, the task demand in Experiments 2 and 3 was of an affective nature, as targets had to be categorized according to their valence (evaluative decision task; EDT). In both the LDT and the EDT, each target stimulus was preceded by a prime stimulus either matching or not matching its affective valence. Using the affective priming paradigm as a procedure, it is possible to examine whether primes are evaluated spontaneously by relying on response times to targets. Trials with prime-target pairs congruent in valence should be relatively faster compared to incongruent trials, and also elicit the same amount or fewer errors.

The chord sequences used never contain the tonic chord in the final position but end without exception on physical major seventh chords. Thus, only the root and therefore also the number of out-of-scale notes of the ending chords is varied, without simultaneously varying the chord type.

In addition to the experimental condition in which chord sequences were used as prime stimuli, a control condition was run in each experiment. In this condition, the final chords of the sequences were presented as single chord primes. This served to ensure that effects found in the experimental condition were not dependent on characteristics of the final chords as standalone entities, but arose from the context they were presented in. In contrast to the study by Sollberger et al. (2003), the present study maintained a constant degree of consonance/dissonance per chord across the categories of positive and negative prime stimuli.

Contrasting results from Experiment 1 with results from Experiment 2 and 3 allows for a discussion of whether spontaneously occurring evaluations of music are dependent on characteristics of the context. More precisely, in Experiment 1, participants carried out a task that diverted the focus from the valence information inherent in the target words, whereas in Experiment 2 and 3 this information was task relevant. Thus, it was possible to investigate whether an affective or valence oriented mindset induced by the primary task affected spontaneous evaluative responses to music, presented as a secondary stimulation.

Experiment 1

In Experiment 1, a nonaffective task (LDT) was employed to examine whether affective priming effects were replicated when the valence dimension was not task relevant. In the field of affective priming research, nonaffective tasks frequently have been used by various researchers (for a summary see Wentura, 2000). So far, heterogeneous results have been reported. Kemp-Wheeler and Hill (1992) obtained evidence for affective priming using a LDT. In addition to nonwords that served as targets, they used words with emotionally aversive or nonemotional connotations as primes and targets. Along a similar line, in a study by Bargh, Chaiken, Raymond, and Hymes (1996), the pronunciation task served as a means to successfully measure affective priming effects. Positive and negative adjectives were used as stimulus material in this study. De Houwer, Hermans, and Eelen (1998) used known words and newly learned nonwords (that were introduced as words from a foreign language) as stimuli but did not observe affective priming effects employing the pronunciation task (for similar results see De Houwer, Hermans, Rothermund, & Wentura, 2002; Klinger, Burton, & Pitts, 2000). In a recent study, Spruyt, De Houwer, Hermans, and Eelen (2007) demonstrated that affective priming effects can be obtained with a nonaffective task if attention is drawn to the valence dimension of the target stimuli.

We expected to observe an affective priming effect in Experiment 1. In line with results from Spruyt et al. (2007), it is conceivable that an affective priming effect is only observed in a particular subgroup of participants: those who report noticing the valence dimension inherent in the category of targets that classify as words in the German lexicon.

Method

Participants. Fifty-two students (29 women, 23 men) from the University of Leipzig participated in the study for either course credit or monetary compensation. Their mean age amounted to 23.15 years (SD = 4.19). All participants were native speakers of German with normal hearing and normal or corrected-to-normal visual acuity. Of the participants, 41.5% were students of psychology in their first year. The data of one female participant had to be excluded due to confounding the answer buttons. Additionally, the data of three further participants (two women, one man) had to be excluded from all analyses due to overall slow responses (determined by the Tukey outlier criterion[2] (Hoaglin, Iglewicz, & Tukey, 1986).

Materials. Forty positive and 40 negative words, the same that were used by Sollberger et al. (2003), served as targets. Based on these target words, 80 pseudowords were constructed by substituting two phonemes of each word with phonemes eliminated from other words in the list, thereby insuring that in total the same phonemes in the 80 original words occurred in the 80 pseudowords (see Appendix).

As primes for the experimental condition, 20 original four-part piano sequences were chosen. These sequences consisted of four triads preceding a final major seventh

[2] Data of participants (Y) that were below (above) the first (third) quartile (Q_1, Q_3) minus (plus) 1.5 times the sample's interquartile range (IQR) [Y < (Q_1 – 1.5*IQR) or Y > (Q_3 + 1.5*IQR)] were discarded.

chord with no specified harmonic progression (a kind of "aborted" functional cadence). They were chosen based on pretest beauty ratings from a total of 84 sequences representing all possible combinations of the seven diatonic steps for the fourth triad root, with the twelve chromatic steps of the final major seventh chord root (see Table 1). The first four chords of the sequences each lasted for 500 ms, whereas the last chord was 1000 ms long, resulting in sequences 3 s in length. The recordings consisted of MIDI-driven (single-key, authentic stereo audio) piano samples with "humanized" volume and timing fluctuations. For humanization, Logic Audio Pro 7.1 was used, with a chance distribution between ± 10 tics (maximum ± 52 ms)—except for the initial chord (maximum +52 ms)—in timing and a chance distribution of the velocity between 64 and 89 (of maximally 127) in the dynamics. As a sampler, we used "Grand Piano" of Soundtrack Pro 2.0.2 by Final Cut/Apple with 16 bit/44.1 kHz stereo audio.

Pretest beauty ratings for the sequences were obtained via the internet from 153 individuals who rated randomly chosen stimulus subsets consisting of 21 sequences. The 10 sequences that were rated as most beautiful on a 5-point scale ranging from 1 ("beautiful") to 5 ("not beautiful") were selected as positive primes (average beauty rating: M = 3.76, SD = 0.81) and 10 sequences rated as less beautiful were selected as negative primes (average beauty rating: M = 2.23, SD = 0.86). The latter were matched regarding the first four chords to the 10 beautiful chord sequences (see Figure 1). This means that the same four chord sequences numbered from I to VII that constituted the beginning of the complete chord sequences were used for both the category of positive and negative prime sequences. In both categories, four-chord-sequence number I was used three times, four-chord-sequence number II once, and four-chord-sequences numbers III, IV, and VI each twice. The beauty ratings of the two final sets differed significantly in a paired t-test, $t(144) = -17.47, p < .001$. They also differed regarding the number of out-of-scale tones contained in the final chords of the sequences. The final chords of positive prime sequences contained fewer out-of-scale tones compared to the final chords of the negative prime sequences, which mostly contained three out-of-scale tones and, in to two cases, two out-of-scale tones (see also Table 2).

As prime stimuli for the control condition, only the last chords of the 20 sequences were used. The last chords of the ten positive sequences were classified as positive primes and the last chords of the ten negative sequences as negative primes. Independent of category, the chords were always major seventh chords only differing in the chroma of the root and the chord structure. Nonetheless, the control condition served as a means to ensure that affective priming effects found in the experimental condition were not attributable to properties of the last chord by itself, but solely to its effects in the context of the preceding four chords.

The presentation was shown on a 15-inch flat screen using Matlab 7.0 (Cogent-Toolbox) to run the experiment. The auditory stimuli were presented binaurally over headphones. Matlab registered judgment responses and latencies, which were captured using the right and left buttons of a four-button response keyboard, beginning 200 ms after the onset of the last chord up to 1800 ms after stimulus presentation.

Procedure. After completing the letter of agreement for participation, participants were tested individually in a dimly lit and sound-attenuated experimental chamber (International Acoustic Company). They were informed that they were about to participate in an experiment designed to study the connection between hearing and reading speed and that their task would be to correctly classify words according to their lexical status, i.e., whether they exist or do not exist in German. Directly after the experiment, participants were questioned about what they had

TABLE 1. Functional Analysis of the First Four Triads Leading to the Seven Diatonic Steps.

No.	Model
I	$T\ S\ D\ T$
II	$T\ T\ S\ S_p$
III	$T\ D\ T_p\ D_p$
IV	$T\ T_p\ D_p\ S$
V	$T\ T_p\ S_p\ D$
VI	$T\ S\ D\ T_p$
VII	$T\ S\ T\ D7$

Note: The last chord (not presented in the table) was always a major seventh chord (D7) or its enharmonic derivates, depending on the imagination of the further progression. Each chord was represented by a letter according to the following schematic: Tonic = T, dominant = D, subdominant = S, parallel = p, counter parallel = c. Majuscule letters represent major triads, minuscule letters represent minor triad, stroked-out letters represent chords missing their root tone.

TABLE 2. Number of Sequences that Contained 0, 1, 2 or 3 Out-of-Scale Tones in its Final Chord Shown for Positive and Negative Prime Sequences.

Number of out-of-scale tones	0	1	2	3
positive prime sequences (10)	4	6	0	0
negative prime sequences (10)	0	0	2	8

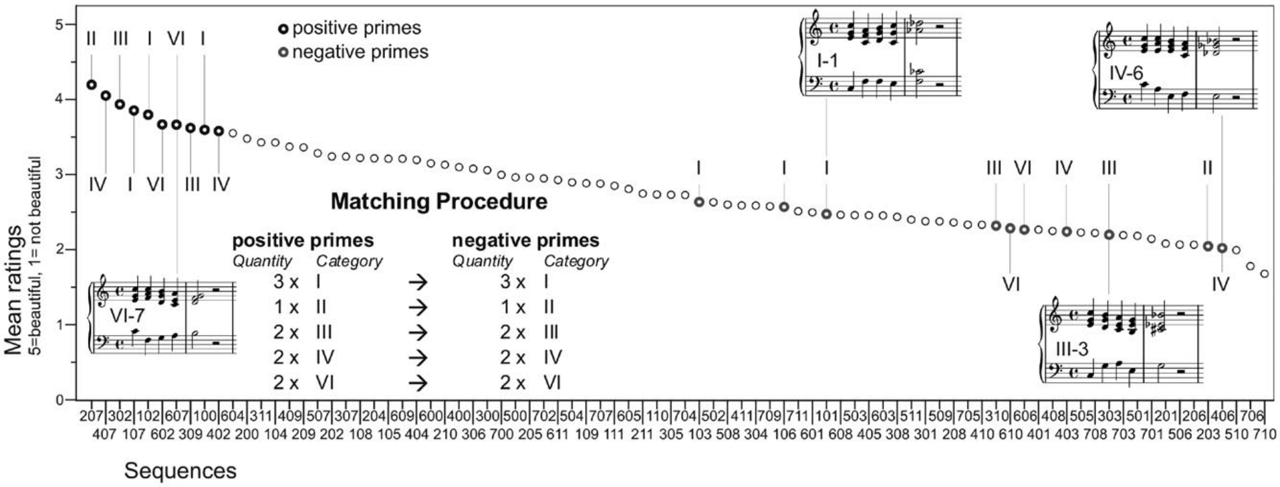

FIGURE 1. Mean pretest beauty ratings of all 84 chord sequences. Ten positive prime sequences (large black circles) and ten negative prime sequences (large gray circles) chosen for the experiments, their respective sequence category number (I, II, III, IV, VI), and music notation for selected sequences.

noticed in regard to the target stimuli in order to assess whether they reported noticing that the target words were either strongly positive or negative. Seventy-five percent of the participants reported noticing the valence dimension without being strongly prompted.

Half of the participants completed the experimental condition first and the control condition subsequently, whereas the order of the conditions was reversed for the other half of the participants. The order in which the conditions were to be completed was counterbalanced across participants. The item sequence was pseudorandomized according to the following constraints: a maximum of three trials of a given lexical status category, a maximum of three trials of a given target valence, and a maximum of three trials of a given prime valence in sequence. The experiment consisted of two blocks per condition: each block included 80 trials, resulting in 160 trials per condition and 320 for the whole experiment. Completion of one block took approximately six minutes in the experimental condition and three minutes in the control condition. Response key assignments (word/pseudoword, pseudoword/word) were counterbalanced across participants.

Before the start of the experiment, participants completed eight practice trials to become familiar with the task. Each trial started simultaneously with a fixation cross shown on the screen and the playback of a chord sequence (experimental condition) or a single chord (control condition). At 250 ms after the onset of the last chord (single chord, respectively), the fixation cross on the screen was replaced by a target word. Targets were presented until participants gave a response or until 1800 ms had elapsed. A blank screen was presented subsequently for a variable uniformly distributed interval (minimum length, 100 ms; maximum length, 500 ms).

For the analysis of response times, all trials with false responses were discarded (2.8% of the data). Additionally, 3.3% of the remaining responses were discarded because they were outliers according to Tukey's outlier criterion (Hoaglin et al., 1986; see also Klauer and Teige-Mocigemba, 2007 for an example); that is, latencies that were below (above) the individual's first (third) quartile minus (plus) 1.5 times the individual's interquartile range were removed. In total, 6.1% of the data were eliminated. Based on these data, individual subject averages were computed. Repeated measures ANOVAs were conducted with the factors Target and Prime with these averages. Similarly, repeated measures ANOVAs were conducted with the factors Target and Prime to analyze error rates. Finally, response time data were subjected to further repeated measures ANOVAs including either the additional factor Valence Detection or Order which refers to the order in which the condition (experimental and control condition) were completed. For the effect central to the main hypothesis, effect sizes are reported as partial η^2 coefficients (Cohen, 1973).

Results

In the experimental condition, the observed pattern in the data was not in accordance with the expectations; that is, trials with positive target words were responded to equally as fast when preceded by a positive prime as when preceded by a negative prime. The same was true for trials with negative target words (see Table 3).

TABLE 3. Mean Response Times (ms) and Standard Deviations for Positive and Negative Words Depending on Prime Category in the LDT (Experiment 1).

Target:	positive		negative	
Prime:	positive	negative	positive	negative
Mean	541.07	542.72	570.30	571.49
SD	70.87	68.53	68.17	71.92

Accordingly, in a repeated measures ANOVA comprising the factors Target and Prime, the interaction between these two factors was not significant, $F(1, 47) = 0.10$, $p = .92$, partial $\eta^2 < .001$.

Overall, response times to positive target words ($M = 541.89$ ms, $SD = 68.97$ ms) were shorter than response times to negative target words ($M = 570.90$ ms, $SD = 68.56$ ms). The main effect of the factor Target was significant, $F(1, 47) = 110.02$, $p < .001$.

The analysis of false responses in the experimental condition revealed that participants produced fewer errors on trials with positive target words ($M = 0.27$, $SD = 0.46$) than on trials with negative target words ($M = 0.89$, $SD = 0.69$). The main effect of the factor Target was significant in a repeated measures ANOVA comprising the factors Target and Prime, $F(1, 47) = 44.57$, $p < .001$. No significant Target x Prime interaction was observed, $F(1, 47) = 0.95$, $p = .33$.

For the experimental condition, the response time data were further analyzed to test whether the group of participants who reported noticing the valence dimension of the target words showed an affective priming effect, as compared to those who did not notice the valence of the words. For this purpose, a repeated measures ANOVA was conducted comprising the factors Target, Prime, and Valence Detection. However, the three-way interaction between these factors was not significant, with $F(1, 43) = 0.19$, $p = .66$; that is, the participants who detected the valence of the words did not differ from those who did not detect it regarding the Target x Prime interaction. Therewith, their data did not show an affective priming effect.

As expected, no affective priming effect was observed in the control condition. This was confirmed with a repeated-measure analysis comprising the factors Target and Prime; there was no significant interaction between these two factors, $F(1, 47) = 0.62$, $p = .44$.

In the experimental condition, participants on average responded more slowly ($M = 581.68$, $SD = 72.24$) than in the control condition ($M = 562.41$ ms, $SD = 66.63$) and produced fewer mistakes ($M = 0.48$, $SD = 0.30$ vs. $M = 0.67$, $SD = 0.48$). Both differences were significant for reaction times, $t(47) = 3.42$, $p = .001$, and error scores, $t(47) = 3.42$, $p = .001$ in paired t-tests.

To test whether response times in the experimental condition were influenced by the order in which the conditions were completed, a repeated measures ANOVA was conducted with the factors Target, Prime, and Order. The interaction between the factors Target and Prime was not influenced by the order in which the conditions occurred, as revealed by a nonsignificant three-way interaction between the factors Target, Prime, and Order, $F(1, 47) < 0.01$, $p = .96$. However, the interaction between the factors Prime and Order was significant in this analysis, $F(1, 47) = 6.81$, $p = .01$. Following up this interaction with separate ANOVAs for the two steps of the factor Order revealed a significant main effect of the factor Prime in the experimental condition for the group of participants who had completed the control condition first, $F(1, 22) = 5.59$, $p = .03$. They responded faster to trials containing positive prime sequences ($M = 569.46$, $SD = 69.44$) than to trials with negative prime sequences ($M = 577.78$, $SD = 67.59$) in the experimental condition. This main effect was not significant for the participants who had completed the experimental condition first, $F(1, 24) = 2.07$, $p = .16$.

Discussion

In the LDT, congruency of valence of prime-target pairs did not affect responses to the targets. Therewith, no affective priming effect was observed. This could mean that the chord sequences were not evaluated spontaneously, or that the method used was insufficient to capture these effects. It cannot be argued that the valence of the words went undetected as 75% of the participants reported noticing that they were presented with words of opposite and extreme polarities. Additionally, in all analyses conducted, main effects of target category were obtained. Trials with positive words rendered faster and more correct responses, compared to trials with negative target words. This indicates that either the affective information contained in the target words was processed (even though this did not result in an affective priming effect) or that the negative words used in the present study occur less frequently in the German language than the positive words, and are therefore more difficult to classify.

Whether or not participants detected the valence dimension did not influence the affective priming effect. In contrast, Spruyt et al. (2007) were able to show the connection between affective priming effects in a nonaffective task and the attention allocated to the affective dimension of the targets. In the present experiment, reports on whether participants strongly focused on the affective content were not obtained. Thus, the measure employed here does not give information on the amount of attention paid to the valence of the words, solely on whether it was noticed at all or not. This might have been crucial for the failure to replicate the results reported by Spruyt et al. (2007).

It was observed that the order in which the conditions were completed moderated the main effect of the factor Prime. Participants who completed the experimental condition subsequent to the control condition responded faster on trials with positive prime sequences than on trials with negative prime sequences in the experimental conditions, independent of the valence of the target words. This effect bears resemblance to the harmonic priming effect that has been reported, for example, by Bigand, Madurell, Tillmann, and Pineau (1999). In harmonic priming studies, faster responses to a target chord have been observed when the target chord functions as the tonic chord in the context of preceding chords, compared to any other chord of the tonal hierarchy (Bigand, Madurell, Tillmann, & Pineau, 1999; Tillmann, Bigand, & Pineau, 1998; Tillmann, Janata, Birk, & Bharucha, 2008). There are two major differences between the present study and the harmonic priming studies mentioned above. First, the stimulus material used in most harmonic priming studies includes the tonic, dominant, or subdominant chord as ending chords, whereas in the present study, neither of these chords is used as an ending chord (i.e., the decisive chords of the present study are farther removed from the tonal centre than in the harmonic priming studies mentioned above). Second, in the harmonic priming studies, participants are instructed to respond to certain characteristics of the musical stimuli. This means that the targets are contained in the musical material. In contrast, the musical stimuli in the present study served exclusively as primes; that is, they were not task relevant. Despite these differences, a similar explanation might apply to both types of priming effects. In the present case, we assume that it cost more effort[3] to process the negative primes than the positive primes. Thus, responses on trials with negative primes were slower due to limited resources for processing the target words. If this effect observed in the present study is interpreted in the light of harmonic priming, the following implications ensue: first, harmonic priming effects may transfer to other modalities (as target word processing was hampered in the present case) and second, they occur even if the decisive harmonic events are peripheral to the tonal center.

It remains unclear why completing the experimental condition subsequent to the control condition is subservient to a type of priming that bears resemblance to harmonic priming, whereas this effect does not occur when the experimental condition is completed first. It is possible to speculate that participants who completed the experimental condition subsequent to the control condition were already familiar with the target words when they completed the experimental condition, as target words were the same in both conditions. Consequently, they might have had more resources to pay attention to the musical stimuli when completing the experimental condition. Thus, as one plausible explanation, the attentiveness with which participants listened to the musical stimuli might be held responsible for the diverging effects. This leads to the hypothesis that harmonic priming is fostered by attentive listening conditions. Further studies are needed in order investigate whether this holds true, and attention is indeed a crucial factor in this matter.

Experiment 2

Experiment 2 directly followed Experiment 1. Again, affective priming effects were studied, this time using a different task demand. Instead of lexical decisions, participants had to perform evaluative decisions by indicating whether target words were either positive or negative. Using the EDT, stable affective priming effects have been obtained in numerous studies (for a summary see, for example, Fazio, 2001).

Method

Participants. After completing Experiment 1, the same participants participated in Experiment 2.

Materials. The same target and prime stimuli that were used in Experiment 1 served as material, with the exception that pseudowords were excluded.

Procedure. Participants were again seated in a dimly lit and sound-attenuated experimental chamber. They were informed that their task was to correctly classify words according to their valence; that is, whether they are positive or negative.

The experiment consisted of two conditions: the experimental condition in which chord progression served as the prime stimuli and the control condition in which only the ending chords served as primes. The order in which the two conditions were to be completed was counterbalanced

[3] Tillmann et al. (2008) observed facilitation for the tonic chord, neither facilitation nor inhibition for the dominant chord, and inhibition for the subdominant chord, compared to a neutral baseline condition in a harmonic priming study. In the present case, the ending chords of the prime sequences were never the tonic or dominant chord. Thus, facilitation or neutrality were less likely to have been effective than inhibition.

across participants. The item sequence was pseudorandomized according to the following constraints: a maximum of three trials of a given target valence and a maximum of three trials of a given prime valence in sequence. The experiment consisted of two blocks per condition; each block included 40 trials, resulting in 80 trials per condition. Completion of one block took approximately six minutes in the experimental condition and three minutes in the control condition. Response key assignments (positive/negative, negative/positive) were counterbalanced across participants. After completing the affective priming task, participants were administered a questionnaire assessing music experience and preferences.

The trial structure was exactly the same as in Experiment 1. This also holds true for the steps taken in preparing the data for analysis. For the analysis of response times, all trials with false responses were discarded (3.5% of the data). Additionally, 3% of the remaining responses were discarded because they were outliers according to Tukey's outlier criterion. In total, 7.5% of the data were eliminated.

The participants were sorted into two groups differing in music expertise according to their experience with music (number of years of music lessons and weekly hours of practice). The group of music experts ($n = 17$) consisted of participants who practiced three or more hours per week or had received music lessons for seven or more years.

Based on the response time data and the error rates, individual subject averages were computed. Repeated measures analyses of variance (ANOVAs) were conducted with the factors Target and Prime with these averages. Similarly, repeated measures ANOVAs were conducted with the factors Target and Prime to analyze error rates. Finally, response time data were subjected to further repeated measures ANOVAs including either the additional factor Order, Field of Study, Gender, or Music Expertise. For effects central to the main hypothesis, effect sizes are reported as partial η^2 coefficients (Cohen, 1973).

Results

In the experimental condition, participants responded faster to trials with prime-target pairs matching in valence than to trials with nonmatching pairs (see Figure 2). This was statistically confirmed in a repeated measures ANOVA comprising the factors Target and Prime, which revealed a significant interaction, $F(1, 47) = 10.10$, $p = .003$, partial $\eta^2 = .18$.

Overall, response times to positive target words (M = 587.88 ms, SD = 77.18 ms) were shorter than response times to negative target words (M = 614.60 ms, SD = 86.98 ms). The main effect of the factor Target was significant, $F(1, 47) = 29.24$, $p < .001$. The analysis of false responses in the experimental condition yielded no significant effects (e.g., Target x Prime: $F(1, 47) = 0.10$, $p = .76$).

As expected, no affective priming effect was observed in the control condition. This was confirmed by the nonsignificant interaction between the factors Target and Prime, $F(1, 47) < 0.001$, $p > .999$. Again, a difference in response times to the two categories of target words was observed. Positive targets were responded to more quickly (M = 581.60 ms, SD = 84.02 ms) than negative targets (M = 601.65 ms, SD = 88.59 ms). This main effect of the factor Target was significant, $F(1, 47) = 13.38$, $p = .001$.

No difference in overall response time between the experimental condition (M = 601.24, SD = 80.42) and the control condition (M = 591.62 ms, SD = 84.22) was observed, $t(47) = 1.52$, $p = .13$. However, participants produced significantly fewer mistakes in the experimental (M = 0.57, SD = 0.49) compared to the control condition (M = 0.79, SD = 0.54), $t(47) = 3.66$, $p = .001$.

To test whether the order in which the conditions were completed influenced response times, a repeated-measure ANOVA was conducted with the factors Target, Prime, and Order. The interaction between the factors Target and Prime was not influenced by the order in which the conditions occurred, as revealed by a nonsignificant three-way interaction between the factors Target, Prime, and Order, $F(1, 46) = 0.51$, $p = .48$. All other effects involving the factor Order were not significant.

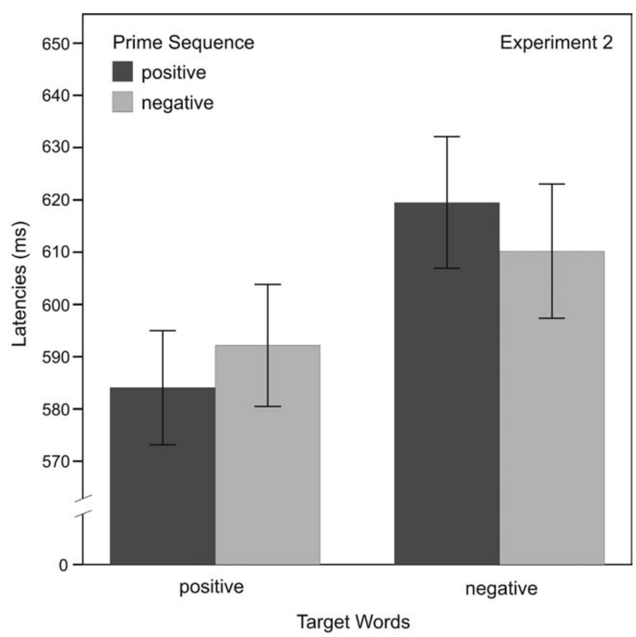

FIGURE 2. Mean response times (in ms) in the EDT (Experiment 2) to positive and negative words as a function of prime valence. Error bars +1,-1 standard errors.

Furthermore, we tested whether the affective priming effect was influenced by whether the participants were students of psychology or students studying another subject. In a repeated measures ANOVA comprising the factors Target, Prime, and Field of Study, the three-way interaction between these three factors was not significant, $F(1, 46) = 0.52$, $p = .82$. Similarly, nonsignificant three-way interaction also emerged when the factor Field of Study was replaced by the factor Gender, $F(1, 46) = 1.03$, $p = .32$, or the factor Music Expertise, $F(1, 46) = 0.20$, $p = .65$. However, as music expertise was of specific interest, this interaction was followed up with separate ANOVAs for each group. The interaction between the factors Target and Prime was significant in the group of laymen, $F(1, 30) = 7.64$, $p = .01$, but not in the group of experts, $F(1, 16) = 2.41$, $p = .14$. It was possible, however, to identify two outliers in the group of experts showing reversed affective priming effects with faster responses to mismatching prime-target pairs than to matching prime-target pairs. Excluding their data from the analysis re-established the significance of the interaction in the subgroup of experts, $F(1, 14) = 11.51$, $p = .004$.

Discussion

In Experiment 2, an evaluative task demand was introduced. Thus, the experimental context in itself was more valence oriented, and therefore more affective in nature than Experiment 1. As a result, a significant affective priming effect was observed in the experimental condition, indicating that spontaneous evaluation of the musical stimuli had occurred. More precisely, trials with prime-target pairs matching in valence received faster responses than trials with prime-target pairs not matching in valence. This means that the chord progressions used as primes were spontaneously evaluated and that this process influenced target processing. Additionally, no affective priming effect was observed in the control condition in which only the ending chords were presented as primes. This shows that the valence of each ending chord depended on its interaction with the context established by the preceding four chords, and not on a quality inherent in the ending chord itself. Thus, contrasting the study by Sollberger et al. (2003), the priming effect observed in the current study did not depend on level of dissonance/consonance or any other characteristic that determines the valence of single chord events.

The observed affective priming effect did not differ between female and male participants or between students majoring in psychology and students majoring in other areas. This shows that the affective priming effect was rather unsusceptible to moderation in the present experiment. Moderating effects of music expertise were analyzed in more depth. An affective priming effect for the group of laymen but not for the group of experts was observed. Terminating the analysis at this step would have led to a discussion about whether laymen spontaneously and affectively react to musical material, whereas experts do not. However, the data were further inspected and after excluding two outliers from the analysis, experts also showed an affective priming effect, rendering it less likely that the above suggestion holds true.

Studying the data of the two participants identified as outliers in more detail revealed that they showed the same unique pattern of music preferences not equalled by any other participants in the sample. Both of them indicated their strong preference for Jazz music, scoring the highest possible value on a scale ranging from one to seven, while they were much less well disposed towards other genres of music.

Reversed affective priming effects—that is, faster responses to prime-target pairs defined as mismatching than to those defined as matching—might emerge if primes intended to be positively evaluated are in fact negatively evaluated and vice versa. It has been shown that stable, as well as very recently acquired preferences (which may of course differ between participants), can be the basis for affective priming effects (Hermans, Baeyens, Lamote, Spruyt, & Eelen, 2005; Hermans, Spruyt, & Eelen, 2003). Therefore, it is possible that the specific pattern of music preferences embraced by the two outliers led them to evaluate the chord progressions in a different manner than other participants.

This interpretation is, to some extent, in line with results of detailed analyses of the pretest data. These analyses showed that music experts' stable genre preference for classical music was a valid predictor for their explicit beauty judgments of the chord sequences. This was not the case for the nonexperts (Müller, Klein, & Jacobsen, 2009; Müller, Klein, & Jacobsen, 2011). However, significant results were exclusively observed for the genre of classical music, not for the genre of jazz music. Therefore, interpretations to that effect have to be considered as speculative. The stimulus material was constructed by focusing on (i) avoiding prejudgments of the stimuli by the experimentators, (ii) the variation of preferably few parameters, and (iii) locating the design of musical allusions in a rather abstract, 19^{th} century style academic field, taking mediocre ecological validity into account. Additionally, the 20 experimental stimuli were chosen by focusing on maximizing predictability of evaluation outcomes. Nonetheless, it seems that the effects of a combination of specific music genre preference and a high amount of music expertise on evaluation outcomes could not be entirely prevented at an individual level. This account is one of the possible explanations for the reversed affective priming effect observed in the two outliers. Further factors that caused reversed priming, especially at the group level, are discussed by Klauer, Teige-Mocigemba, and Spruyt (2009).

All participants who took part in Experiment 2 had also participated in Experiment 1. Thus, it is disputable whether the significant affective priming effect observed in Experiment 2 was indeed due to the change in task; the possibility cannot be excluded that it was due to the fact that the participants were already familiar with the stimulus material after having completed Experiment 1. To clarify this issue, Experiment 3 was conducted.

Experiment 3

Experiment 3 was a replication of Experiment 2, with participants that had not participated in Experiment 1 and 2. Additionally, participants explicitly rated the musical stimuli with respect to their beauty and valence, subsequent to the completion of the EDT. This additional task was introduced to assure conformity with the results of the pretest. Furthermore, musical stimuli were classified according to their beauty in the pretest, whereas in the experiments, valence was the central concept in all tasks. Thus, in the case of the present stimulus material, it was necessary to control whether ratings of beauty and valence overlapped sufficiently to be treated as interchangeable category labels. A scale spanning from very beautiful to not beautiful was chosen for the explicit rating. The same scale was used in the pretest. This scale was chosen because it has recently been confirmed in a questionnaire study that the adjective "beautiful" resides at the core of the conceptual structure of music aesthetics (Istók et al., 2009). This adjective therefore has been found to represent the optimal linguistic device for expressing the aesthetic value of music (Istók et al., 2009). As a second scale, the valence dimension spanning from positive to negative was chosen for the explicit rating because of its importance for the EDT. Furthermore, participants were asked for their personal hypothesis concerning the aim of the experiment. In line with the results of Sollberger et al. (2003), it is expected that the affective priming effect is independent of whether participants make a correct guess about the experiment's hypothesis or not.

Method

Participants. Forty students (24 women, 16 men) from the University of Leipzig participated in the study for either course credit or monetary compensation. Their mean age was 21.93 years ($SD = 4.35$). All participants were native speakers of German with normal hearing and normal or corrected-to-normal visual acuity. Of the participants, 70% were students of psychology in their first year. The data of one male participant had to be excluded from all analyses due to overall slow responses (determined by the Tukey outlier criterion). Hence, all analyses reported in the following are based on the data of the 39 remaining participants.

Materials. The same target and prime stimuli that were used in Experiment 2 served as material.

Procedure. The procedure was identical to that of Experiment 2. Additionally, explicit ratings of the chord progressions were obtained following the completion of the EDT. Participants judged either the valence or the beauty of the chord progressions first. The order in which the tasks were to be completed was counterbalanced across participants. Five-point Likert scales were used for this purpose with the endpoints labelled "not beautiful" and "very beautiful" in the explicit beauty rating task, and "positive" and "negative" in the explicit valence rating task.

Subsequent to completing all tasks and filling in the questionnaires, participants were asked for their personal hypothesis regarding the objective of the experiment. Seventeen participants made a good guess at the hypothesis; that is, they correctly guessed that the affective match or mismatch between musical prime stimuli and target words was supposed to affect the processing of the target words. These 17 participants were henceforth considered as the group of participants who correctly guessed the hypothesis, even though only eight of them guessed that the main focus of the EDT lay in affecting response times. Error rates were mostly named as the dependent measure of interest or the perception of the words in general (without the respondents being sure how this could be measured). Nineteen participants claimed to have no idea, or put forth incomplete or false hypotheses. For three participants, no data were acquired on this measure.

Following the same procedure as described in Experiment 2, participants were sorted into two groups differing in music experience. The group of participants high in musical experience was comprised of 17 individuals and the group of participants with less experience 22 individuals.

For the analysis of response times, all trials with false responses were discarded (3.3% of the data). Additionally, 4.0% of the remaining responses were discarded because they were either too fast or too slow according to Tukey's outlier criterion. In total, 7.3% of the data were eliminated. Individual subject averages were computed based on the remaining data. Repeated measures analyses of variance (ANOVAs) were conducted as in Experiment 2. Additionally, repeated measures ANOVAs were computed with the factors Target, Prime, and Hypothesis.

Results

In the experimental condition, participants responded faster to trials with prime-target pairs matching in valence than to trials with nonmatching pairs (see Figure 3). This was statistically confirmed in a repeated measures ANOVA comprising the factors Target and Prime by the significant interaction between these two factors, $F(1, 38) = 6.46$, $p = .02$, partial $\eta^2 = .15$.

The main effects in this analysis were not significant. Furthermore, the analysis of false responses in the experimental condition yielded no significant effects (e.g., Target x Prime: $F(1, 38) = 1.11$, $p = .30$).

As expected, no affective priming effect was observed in the control condition. This was confirmed by the nonsignificant interaction between the factors Target and Prime, $F(1, 38) = 0.53$, $p = .47$. A difference in response times to the two categories of target words was observed. Positive targets were responded to more quickly (M = 572.27 ms, SD = 80.22 ms) than negative targets (M = 591.49 ms, SD = 85.98 ms). This main effect of the factor target was significant, $F(1, 38) = 14.00$, $p = .001$.

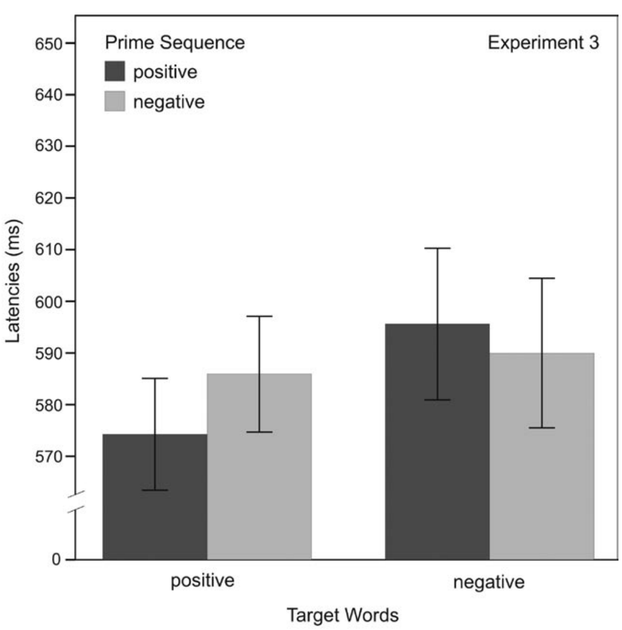

FIGURE 3. Mean response times (in ms) in the EDT (Experiment 3) to positive and negative words as a function of prime valence. Error bars +1,-1 standard errors.

No difference in overall response time and the rate of false responses between the experimental condition (M = 586.45 ms, SD = 76.67 ms; M = 0.63, SD = 0.58) and the control condition (M = 581.88 ms, SD = 81.59 ms; M = 0.67, SD = 0.63) was observed, $t(38) = 0.43$, $p = .67$ and $t(38) = 0.54$, $p = .59$.

To test whether the order in which the conditions were completed had an influence on response times, a repeated-measure ANOVA was conducted with the factors Target, Prime, and Order. The interaction between the factors Target and Prime was not influenced by the order in which the conditions occurred, as revealed by a nonsignificant three-way interaction between the factors Target, Prime, and Order, $F(1, 38) = 0.04$, $p = .85$. All other effects involving the factor Order were also not significant.

Furthermore, we tested whether the affective priming effect was influenced by whether the participants were students of psychology or students studying another subject. In a repeated measures ANOVA comprising the factors Target, Prime, and Field of Study, the three-way interaction was not significant, $F(1, 37) = 1.39$, $p = .25$. Similarly, a nonsignificant three-way interaction also emerged when the factor Field of Study was replaced by the factor Gender, $F(1, 37) = 1.36$, $p = .25$, or Music Expertise, $F(1, 37) < 0.01$, $p = .98$.

To test whether guessing the hypothesis correctly or not affected response times in the experimental condition, a repeated-measure ANOVA was conducted with the factors Target, Prime, and Hypothesis. The three-way interaction was not significant, $F(1, 34) = 1.76$, $p = .19$. The same was true for the two-way interaction between the factors Prime and Hypothesis, $F(1, 34) = 3.93$, $p = .06$. However, the two-way interaction between the factors Target and Hypothesis was significant, $F(1, 34) = 5.82$, $p = .02$.

In the analysis of error rates, the interaction between the factors Target, Prime, and Hypothesis, the interaction between the factors Target and Hypothesis, and the interaction between the factors Prime and Hypothesis were not significant, $F(1, 34) = 1.10$, $p = .30$, $F(1, 34) = 0.01$, $p = .93$, and $F(1, 34) = 1.94$, $p = .17$, respectively. The same was true for the analysis of response times (all Fs < 1.00) and error rates in the control condition (all Fs < 1.30).

The explicit valence and beauty ratings of the participants in Experiment 3 were compared with the beauty ratings obtained from the pretest sample (see Figure 4). Based on the beauty ratings of the experimental sample, the resulting categorization of positive and negative primes was the same as the original categorization of chord progressions. Based on the valence ratings of the

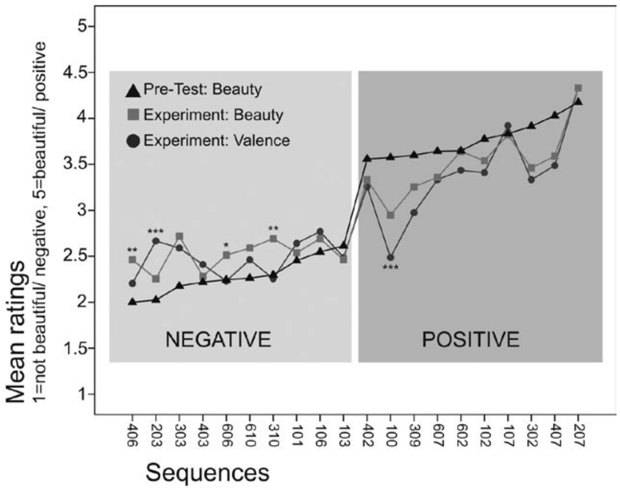

FIGURE 4. Mean beauty and valence ratings for all positive (dark gray box) and negative (light gray box) prime sequences. Mean beauty ratings of the pretest sample are depicted as black triangles, mean beauty ratings of the experimental sample as light gray squares, and mean valence ratings of the experimental sample as dark gray circles. Significant differences between the mean beauty and the mean valence ratings of the experimental sample are marked as follows: * $p < .10$, ** $p < .05$, *** $p < .01$.

experimental sample, the categorization that resulted was to a large extent identical to the original categorization. This means that of the ten chord progressions that were included in the category of negative prime stimuli, nine were categorized in a manner corresponding to the original categorization. The same was true for the category of positive prime stimuli.

The correspondence between pretest and experimental ratings was further quantified by correlation analyses. Pretest ratings correlated significantly with beauty ratings, $r(18) = .93$, $p < .001$, and valence ratings, $r(18) = .86$, $p < .001$, of the experimental sample. Similarly, the beauty and the valence rating of the experimental sample were also significantly correlated, $r(18) = .94$, $p < .001$. Mean rating scores for both sets are shown in Table 4.

Discussion

In line with the results of Experiment 2, the expected affective priming effect was observed in Experiment 3. This shows that the affective priming effect observed in the EDT is also replicated if participants have not heard the musical stimulus material before and completed the LDT as a preceding task. Thus, the effect does not depend on prior familiarity with the stimulus material.

Furthermore, explicit ratings of the stimuli were obtained. Participants rated the beauty as well as the valence of the prime stimuli subsequent to the EDT.

TABLE 4. Mean Rating Values and Standard Deviations for Both Sets of Prime Stimuli.

	Pretest sample	Experimental sample	
	Beauty rating	Beauty rating	Valence rating
Positive set *Mean (SD)*	3.72 (0.20)	3.53 (0.37)	3.40 (0.49)
Negative set *Mean (SD)*	2.23 (0.21)	2.52 (0.16)	2.47 (0.20)

Note: Stimuli were rated on 5-point Likert scales ranging from 1 = "not beautiful"/ "negative" to 5 = "beautiful"/"positive."

Significant differences between beauty and valence ratings were observed in only four out of 20 stimuli (see Figure 4). This finding, combined with the high correlation coefficient observed between beauty and valence scores, leads to this assumption: that in the case of the present stimulus material, the concepts of beauty and valence are closely related enough to be treated as interchangeable category labels in the scope of this study. Furthermore, the explicit ratings of the experimental sample were in line with the ratings of the pretest sample.

Following Sollberger et al. (2003), we tested whether the participants that had made a good guess at the hypothesis of the experiment differed from those who had made either no guess or an inaccurate guess. In line with the results of Sollberger et al. (2003), the affective priming effect was not significantly moderated by hypothesis group.

General Discussion

In the scope of the present study, we examined evaluative processes occurring spontaneously in response to musical stimulus material. Three experiments were conducted using the affective priming paradigm as a means to study these processes implicitly. The experiments differed according to how strongly attention was drawn to the valence of the target words that were used, in order to measure spontaneous evaluation of the musical stimuli that served as primes. Experiment 1 consisted of an LDT, whereas in Experiments 2 and 3, target words had to be categorized according to their valence (EDT). The task employed in Experiment 1 was therefore less evaluative in nature than the task employed in Experiments 2 and 3.

By obtaining affective priming effects in Experiments 2 and 3, we were able to demonstrate the occurrence of unprompted evaluation processes in response to the ending chords of chord progressions. Observing no similar priming effect in Experiment 1 raises this question: is musical material evaluated spontaneously only if it is presented in a context that classifies as evaluative? (e.g., the context created by Experiment 2 with an inherently high saliency of the valence dimension). Or, alternatively: is musical material invariably evaluated independent of context variables? If this is the case, the use of an LDT in Experiment 1 would have failed to capture this process. The former account is in line with results from a study conducted by Spruyt et al. (2007), who found that "automatic stimulus processing depends on the extent to which the experimental context, as a whole, encourages participants to assign attention to the affective stimulus dimension" (p. 48). In contrast, the latter account is in line with Deutsch and Gawronski (2009), who caution against prematurely interpreting differences between priming effects as authentic modulations of the underlying evaluation process. Based on the current data it cannot be terminally determined which of the two alternative accounts holds true.

Koelsch et al. (2004) have concluded that musical stimuli, similar to linguistic stimuli, are subject to automatic semantic analyses. Thus, it is important to discuss and delineate the affective priming effects obtained in the present study in light of these results. Compared to affective priming effects, semantic priming effects are replicated much more consistently with LDTs and pronunciation tasks (Storbeck & Robinson, 2004). According to Storbeck and Robinson, observing affective priming effects with an LDT might be evidence that additional semantic information contained in the musical material distinguishes the categories of positive and negative stimuli from each other. For example, in a study that uses only words as primes and targets, this would be the case if all negative words referred to unpopular animals (e.g., spider, snake) and all positive words to flowers (e.g., rose, lily). In such a case, the priming effect observed could not be labelled "affective," as it would probably be less driven by the affective than by the additional semantic information contained by the stimuli. In the present study, no priming effect was observed in the LDT. This hints at the fact that the musical stimuli used in the present study lack additional semantic information that varies between categories. As a consequence, it was only possible to assign category membership based on valence information. Thus, affective information was exclusively responsible for the affective priming effects observed in Experiments 2 and 3 with the EDT. This strengthens the claim that we have truly captured spontaneous evaluative processes in the present experiments and not the residues of automatic semantic analyses.

Furthermore, Storbeck and Robinson (2004) compared affective and semantic priming effects to contribute to the debate on whether the claim of affective primacy (Zajonc, 1980) is to be upheld, or whether semantic analyses are the precondition for affective processing. They observed more robust semantic than affective priming effects in four out of five experiments. Affective priming effects emerged only in their third experiment, in which the stimulus material was reduced to a set of words that contained only affective without additional semantic information. The authors concluded that "people may only retrieve affective associations when initial analyses in terms of semantic category membership yield minimal information" (p. 86). This is a very interesting finding regarding music, which, compared to language in general, is less rich in clear-cut semantic information; therefore, automatic attempts at semantically categorizing musical material very often results in "minimal information." Combining this insight with one contributed by Koelsch et al. (2004), it becomes conceivable that semantic analyses of musical stimuli take place but remain relatively fruitless in many cases. This, in turn, leads to an increased retrieval of affective associations. Thus, the relative lack of semantic content in music might be a crucial factor motivating the affective processing of music. The power of music to evoke emotions is a well-documented phenomenon, and in recent years the importance of studying the mechanisms responsible for eliciting emotions in response to music has been emphasized (for example, see Juslin & Västfjäll, 2008). The mechanism outlined here might add another piece to the puzzle in explaining the strong link between music and emotion.

With the present study, we extend findings previously reported by Sollberger et al. (2003) by showing that spontaneous evaluation transcends the level of single sound events, and also applies to situations in which a certain polarity emerges from the relations between multiple sound events. It is especially interesting that the musical chords used in the present study had neither a specific polarity of their own nor gained it through effects of contrast; that is, a stimulus evaluated as being exceptionally beautiful was not followed by one rated as neutral, which would have rendered the latter slightly negative in comparison. Also, the ending chords did not vary with respect to their adherence to music stylistic rules: the relation of the ending chord to the preceding four chords was functionally valid in a number of historical Western music styles (mostly Baroque at least), and composed correctly after Baroque piano voicing principles (basso continuo) for all sequences. The overall impression of the stimuli, being in the field of 18^{th}–19^{th} century Western artistic music, was not too closely related to any specific music style (as for example the idiom of Wiener Klassik), such that specific expectations of progressions and their fulfilment or disappointment (up to the feeling of stylistic "violations") were avoided and left to individual preference and imagination. Moreover, the amount of the ending effect was comparable, as none of the sequences ended on a functionally closing chord (e.g., tonic).

Furthermore, even though their behavioral data did not capture local changes in affective responses to altered chords, Steinbeis et al. (2006) reported ERP data pointing to the fact that these changes occur. This is confirmed by the present results, which provide clear evidence for immediate affective evaluative processes occurring locally and unprompted during music reception.

Conclusion

Music was presented as a secondary stimulus in the present study, thereby recreating a very common listening situation: that of music playing in the background while participants were occupied with another primary task. It was demonstrated that music is evaluated even in the absence of the demand for explicit judgments. This is a finding of fundamental importance to the discussion of larger scale emotions felt in reaction to music. There was also evidence that raises the question of whether spontaneous evaluation processes are sensitive to context manipulations. The discussion of the current results with reference to harmonic and semantic priming studies presents a starting point for a more comprehensive approach to future research on implicit music processing.

Author Note

The authors thank Sara Bergman for proofreading and three anonymous reviewers for very helpful comments on an earlier version of this article. This work was supported by the NEST (New and Emerging Science and Technology) program of the European Commission (FP6-2004-NEST-PATH-028570). The musical stimuli used in the study can be listened to on the following web-page: www.uni-leipzig.de/~brain/stimuli, Username: music, Password: perception.

Correspondence concerning this article should be addressed to Mira Müller, Department of Psychology I, University of Leipzig, Seeburgstraße 14-20, D-04103 Leipzig, Germany. E-mail: mira.mueller@uni-leipzig.de

References

Altenmüller, E., Schürmann, K., Lim, V. K., & Parlitz, D. (2002). Hits to the left, flops to the right: Different emotions during listening to music are reflected in cortical lateralisation patterns. *Neuropsychologia, 40*, 2242–2256.

Bargh, J. A., Chaiken, S., Raymond, P., & Hymes, C. (1996). The automatic evaluation effect: Unconditional automatic attitude activation with a pronunciation task. *Journal of Experimental Social Psychology, 32*, 104–128.

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist, 54*, 462–479.

Bigand, E., Madurell, F., Tillmann, B., & Pineau, M. (1999). Effect of global structure and temporal organization on chord processing. *Journal of Experimental Psychology-Human Perception and Performance, 25*, 184–197.

Cohen, J. (1973). Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educational and Psychological Measurement, 33*, 107–112.

De Houwer, J., Hermans, D., & Eelen, P. (1998). Affective and identity priming with episodically associated stimuli. *Cognition and Emotion, 12*, 145–169.

De Houwer, J., Hermans, D., Rothermund, K., & Wentura, D. (2002). Affective priming of semantic categorisation responses. *Cognition and Emotion, 16*, 643–666.

Deutsch, R., & Gawronski, B. (2009). When the method makes a difference: Antagonistic effects on "automatic evaluations" as a function of task characteristics of the measure. *Journal of Experimental Social Psychology, 45*, 101–114.

Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition and Emotion, 15*, 115–141.

Hermans, D., Baeyens, F., Lamote, S., Spruyt, A., & Eelen, P. (2005). Affective priming as an indirect measure of food preferences acquired through odor conditioning. *Experimental Psychology, 52*, 180–186.

Hermans, D., Spruyt, A., & Eelen, P. (2003). Automatic affective priming of recently acquired stimulus valence: Priming at SOA 300 but not at SOA 1000. *Cognition and Emotion, 17*, 83–99.

Hoaglin, D. C., Iglewicz, B., & Tukey, J. W. (1986). Performance of some resistant rules for outlier labeling. *Journal of the American Statistical Association, 81*, 991–999.

Istók, E., Brattico, E., Jacobsen, T., Krohn, K., Müller, M., & Tervaniemi, M. (2009). Aesthetic responses to music: A questionnaire study. *Musicae Scientiae, 13*, 183–206.

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences, 31*, 559–575.

Kemp-Wheeler, S. M., & Hill, A. B. (1992). Semantic and emotional priming below objective detection threshold. *Cognition and Emotion, 6*, 113–128.

Klauer, K. C., & Teige-Mocigemba, S. (2007). Controllability and resource dependence in automatic evaluation. *Journal of Experimental Social Psychology, 43*, 648–655.

Klauer, K. C., Teige-Mocigemba, S., & Spruyt, A. (2009). Contrast effects in spontaneous evaluations: A psychophysical account. *Journal of Personality and Social Psychology, 96*, 265–287.

Klinger, M. R., Burton, P. C., & Pitts, G. S. (2000). Mechanisms of unconscious priming: I. Response competition, not spreading activation. *Journal of Experimental Psychology-Learning Memory and Cognition, 26*, 441–455.

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Friederici, A. D. (2004). Music, language and meaning: Brain signatures of semantic processing. *Nature Neuroscience, 7*, 302–307.

Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago, IL: University of Chicago Press.

Müller, M., Höfel, L., Brattico, E., & Jacobsen, T. (2010). Aesthetic judgments of music in experts and laypersons–An ERP study. *International Journal of Psychophysiology, 76*, 40–51.

Müller, M., Klein, J., & Jacobsen, T. (2009, February). *Aesthetic evaluation of chord progressions in piano music*. Poster presented at the BrainTuning Workshop: Music, Emotions, and Brain Plasticity, Helsinki, Finland. doi: 10.3389/conf.neuro.09.2009.02.025

Müller, M., Klein, J., & Jacobsen, T. (2011). *Musical experience and genre preferences run deep: A study of differential effects on aesthetic responses to chord progressions in piano music*. Manuscript submitted for publication.

Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness - Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology, 64*, 723–739.

Orr, M. G., & Ohlsson, S. (2005). Relationship between complexity and liking as a function of expertise. *Music Perception, 22*, 583–611.

Ortony, A., Clore, G., & Collins, A. (1988). *The cognitive structures of emotions*. Cambridge, MA: Cambridge University Press.

Peretz, I., Gaudreau, D., & Bonnel, A. M. (1998). Exposure effects on music preference and recognition. *Memory and Cognition, 26*, 884–902.

Scherer, K. R. (2004). Which emotions can be induced by music? What are the underlying mechanisms? And how can we measure them? *Journal of New Music Research, 33*, 239–251.

Scherer, K. R., & Zentner, M. (2008). Music evoked emotions are different – More often aesthetic than utilitarian. *Behavioral and Brain Sciences, 31*, 595–596.

Sloboda, J. A., & O'Neill, S. A. (2001). Emotions in everyday listening to music. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion* (pp. 415–429). New York: Oxford University Press.

Sollberger, B., Reber, R., & Eckstein, D. (2003). Musical chords as affective priming context in a word-evaluation task. *Music Perception, 20*, 263–282.

Spruyt, A., De Houwer, J., Hermans, D., & Eelen, P. (2007). Affective priming of nonaffective semantic categorization responses. *Experimental Psychology, 54*, 44–53.

Steinbeis, N., Koelsch, S., & Sloboda, J. A. (2006). The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. *Journal of Cognitive Neuroscience, 18*, 1380–1393.

Storbeck, J., & Robinson, M. D. (2004). Preferences and inferences in encoding visual objects: A systematic comparison of semantic and affective priming. *Personality and Social Psychology Bulletin, 30*, 81–93.

Tillmann, B., Bigand, E., & Pineau, M. (1998). Effects of global and local contexts on harmonic expectancy. *Music Perception, 16*, 99–117.

Tillmann, B., Janata, P., Birk, J., & Bharucha, J. J. (2008). Tonal centers and expectancy: Facilitation or inhibition of chords at the top of the harmonic hierarchy? *Journal of Experimental Psychology: Human Perception and Performance, 34*, 1031–1043.

Wentura, D. (2000). Dissociative affective and associative priming effects in the lexical decision task: Yes versus no responses to word targets reveal evaluative judgment tendencies. *Journal of Experimental Psychology: Learning Memory and Cognition, 26*, 456–469.

Wittenbrink, B. (2007). Measuring attitudes through priming. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitude: Procedures and controversies* (pp. 17–58). New York: Guilford Press.

Zajonc, R. B. (1980). Feeling and thinking – Preferences need no inferences. *American Psychologist, 35*, 151–175.

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion, 8*, 494–521.

APPENDIX. Target Stimuli

Positive				**Negative**			
Words		*Pseudo*		*Words*		*Pseudo*	
Liebe	Oase	**Kiere**	Eane	**Mörder**	Sklave	**Mirler**	Sklede
Freude	Wärme	Draude	Börme	Gewalt	Gier	Hezalt	Wiet
Glück	Charme	Grack	Chilme	Folter	Akne	Goller	Afna
Lachen	Blume	Nechen	Gläme	Inzest	Rache	Anzert	Teche
Urlaub	Herz	Arleub	Ferv	Hass	Kotzen	Hurs	Watzen
Spaß	Party	Spüb	Surty	Unfall	Übel	Inkall	Üpal
Sonne	Erfolg	Bonfe	Arfoll	Qual	Hölle	Quor	Gölla
Humor	Spiel	Heror	Spoeh	Horror	Panzer	Morrur	Halzer
Lob	Jubel	Peb	Lubol	Elend	Geisel	Eluns	Jeusel
Lust	Trost	Sugt	Tresk	Seuche	Betrug	Leucho	Setreg
Kuss	Kraft	Kubt	Krarr	Verrat	Armut	Vessat	Asmuß
Ferien	Fest	Letien	Fosp	Leid	Ekel	Keed	Eful
Reise	Gewinn	Reule	Wesinn	Unheil	Satan	Inheis	Gadan
Musik	Meer	Futik	Rier	Strafe	Frust	Strile	Framt
Genuss	Licht	Meness	Gucht	Not	Sucht	Som	Zicht
Natur	Geburt	Tatul	Genunt	Geiz	Pein	Seuz	Heit
Schatz	Reife	Schagl	Seufe	Zwang	Kot	Zwont	Züt
Dank	Ruhe	Dülk	Rane	Trauma	Ruin	Prauna	Wain
Wunder	Gefühl	Kubder	Gefehn	Opfer	Feind	Uprer	Reint
Leben	Perle	Rebun	Setle	Ekzem	Lüge	Efzim	Foge