Labelling Musical Harmonies in Emotional Terms

Linus Backström (202004875)

Faculty of Arts, Aarhus University

Cognition and Communication

Bachelor of Cognitive Science

February 11, 2021

Abstract (200-300 words)

Music has been used throughout history to convey emotions, but what are the underlying features that contribute to music perception? Previous research suggests that emotions in music may relate to the listener's anticipation of how the music progresses. The focus of this study was to investigate perceived emotions in musical harmony. A repeated measures experiment was conducted to measure how different variations of simple, computer generated two-chord progressions affect ratings of valence and arousal, and whether coherent structures would emerge from these ratings. The musical background of participants was assessed and used to identify potential differences in ratings between musicians and nonmusicians. The study found a significant negative correlation between ratings of valence and the number of non-diatonic notes per progression. The study also found that musicians were more consistent at rating chord progressions in terms of arousal and valence, as measured by ICC scores.

Keywords: arousal, valence, music perception, harmony

The language of music

Consider the different aspects of human language for a moment. Although the actual words we speak carry a fair amount of information, we also depend on a variety of non-verbal cues to effectively communicate with each other. The speaker's tone, emphasis on certain words, accompanying gestures and facial expressions all affect how a message is interpreted. In a similar manner our perceptions of a musical piece are dependant not only on the chords and harmonies, but also the tempo, dynamics, timbre, rhythm, etc. In the literature these cues that exist outside of specific tonal systems (such as the 12-tone system in Western music) have been referred to as *psychophysical dimensions of music,* and they are defined as properties of sounds that the listener can perceive independent of culture and musical experience (Balkwill & Thompson, 1999).

In spoken language, a strong and clear message involves verbal and non-verbal cues that are congruent with each other. An example of this could be when a person looks down with a quivering lower lip and quietly says “I am feeling sad”. The musical equivalent could arguably be when a chord or harmony typically associated with sadness, such as the minor chord in Western music, is played softly (“piano” in musical terms) and at a lower tempo (“lento”), perhaps on a violin or piano. Both of these examples describe situations where the listener receives many cues that point towards the intended emotion. On the other hand, if a speaker announces “I am feeling sad!” in a cheerful voice while smiling and dancing, the listener would be unlikely to take the words at face value, as they are incongruent with the other cues. For the musical equivalent of a mixed message, imagine minor chords being played on a banjo at a high tempo.

Brain imaging studies (Koelsch & Friederici, 2003; Koelsch et al., 2004) using electroencephalography (EEG) and magnetoencephalography (MEG) have found that when a listener hears a harmonically incongruous chord during a musical sequence, a brain response called the “early right anterior negativity” is elicited. Koelsch and colleagues traced the source of the response to a component of Broca’s region known as the inferior frontal cortex, which is an area frequently associated with the processing of syntax in language (Vuust & Kringelbach, 2010).

To finish off the comparison between our spoken language and music, consider the possibility that just as English-speaking people have a concept of sadness that is activated when the three arbitrary sounds “S”, “A” and “D” are spoken in succession to form the word “sad”, a similar mechanism exists for how we interpret music. People who have been regularly exposed to Western tonal music have learned that when the three notes “D”, “F” and “A” are played in succession (or simultaneously) they form the chord “D-minor” – a chord often associated with sadness. This proposed relationship between letters and notes, words and chords, and even phrases and harmonies, is not perfect by any means and requires more research. It can however be used as an analogy for conceptualizing the underlying mechanisms of how music is interpreted at its most basic level.

The impact of prior knowledge and culture on music perception

In the field of psychology researchers have long debated over the role of nature versus nurture. Those who study music engage in a similar debate about which aspects of music perception are universal and which have been taught by the musical culture that the listener has been exposed to (Balkwill & Thompson, 1999; McDERMOTT & Hauser, 2005; Peretz, 2006; Trehub, 2003). Balkwill & Thompson (1999) designed a study around the idea that when the listener is unfamiliar with a particular tonal system, they focus on psychophysical dimensions to recognize musically expressed emotions. The authors found that listeners were indeed able to recognize intended emotions in unfamiliar tonal systems (their study used Hindustani ragas) and that they used psychophysical dimensions (tempo, timbre, and complexity) to achieve this.

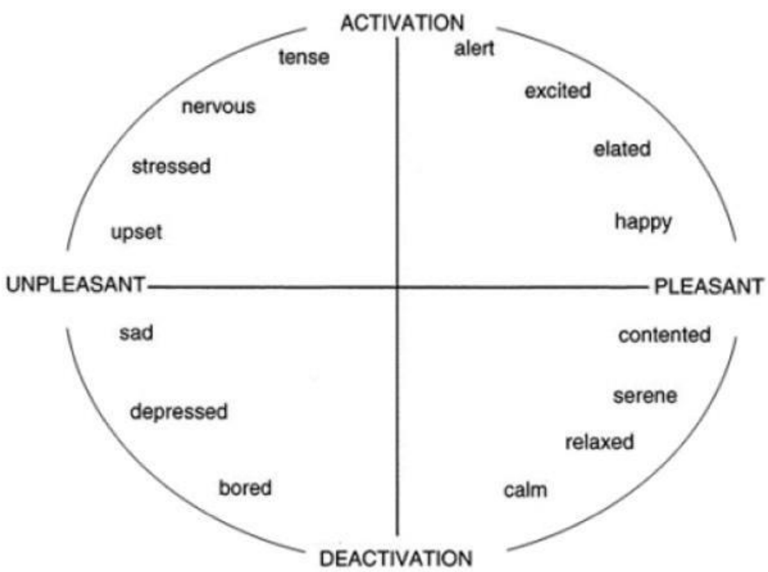
In some sense, the present study takes the opposite approach, as it studies listeners who are familiar with the tonal system and their ability to perceive emotions when the music is stripped of these psychophysical dimensions. In other words, this study recognizes that there are several aspects of music that contribute to the listeners perceptions and attempts to isolate one of these aspects, i.e. harmony, to find out whether significant differences in perceived emotions can be induced merely by varying the harmony, while keeping the other aspects constant.

It is also worth mentioning that while an important question, this study was not designed to answer the question of whether differences in perceived emotions in music exist due to cultural background, previous experience, or something else entirely. Nevertheless, the results provide some insight into how musical training affects perceived emotions.

The two-dimensional circumplex model of emotions

Music and emotion research has generally seen the use of two different models of emotions in music, and these are the discrete and dimensional models (Eerola & Vuoskoski, 2011). The discrete model places emotions into different categories, while the two-dimensional model defines them using varying degrees of arousal and valence. There is also support from neurological studies that the processes involved in these two ways of assessing emotions are relatively independent of each other (Gosselin, 2005; Khalfa et al., 2008).

Figure 1. The two-dimensional circumplex model



Note: Figure is from Posner et al. (2005)

According to the two-dimensional circumplex model (Posner et al., 2005), all emotions are rooted in two independently functioning neurophysiological systems: an activation-deactivation continuum (known as arousal), and a pleasure-displeasure continuum (known as valence). As mentioned, the two-dimensional model is often used in music and emotion research (Gomez & Danuser, 2004; Witvliet & Vrana, 2007), and is also supported by findings showing that emotions are relatively quickly and easily evaluated in terms of valence and arousal (Vieillard et al., 2008). The model has on the other hand been critiqued for not being able to differentiate between emotions that lie close to each other in the valence-activation space, e.g. anger and fear (Tellegen et al., 1999). Others have pointed out that the model is poor at capturing some of the finer emotions that music is capable of expressing (Bigand et al., 2005; Collier, 2007; Ilie & Thompson, 2006). However, since the musical stimuli used in the present study was reduced to simple two-chord progressions, capturing finer emotions was not required of the model. The advantages of the two dimensional-model – primarily the relative ease and speed at which participants were shown to be able to rate emotions using arousal and valence (Vieillard et al., 2008), and the continuous data gathered from these ratings – were therefore judged to outweigh the limitations of the model.

Anticipation and prediction

When researching music and emotions it quickly becomes clear that two related concepts, anticipation and prediction, are viewed by leading experts as being essential for gaining a deeper understanding about how music and emotions are connected. Vuust & Kringelbach (2010) describe the connection as follows: "[…] an emotion is induced in a listener because a specific feature of the music violates, delays, or confirms the listener's anticipation of the continuation of the music" (p. 170). This implies that the emotions associated with music may differ depending on whether the musical sequence is unpredictable or simply continues as expected. In the book "Sweet Anticipation", Huron (2006) identifies the "prediction effect", where positively valenced emotions are evoked when listeners successfully predict future musical events. Huron goes on to describe different ways by which this can be achieved, with the most relevant one being "schematic predictability", which occurs when the music matches the listeners existing schemas. These schemas will vary between cultures, but the important schemas for the present study are those which relate to the familiar harmonies of Western music.

Diatonic chords are chords that naturally occur within a certain key signature, and virtually any composition in Western music consists mostly, often entirely, of diatonic chords – whether it is classical music, popular music, electronic music, or nearly any other genre. Exceptions to this rule are sure to be found (e.g. in modal jazz), but even so most would agree that exposure to Western music teaches the ear to expect diatonic harmonies. This long term learning of musical structures has been called "culture-dependent statistical learning" in the literature (Vuust & Kringelbach, 2010). Finally, it has been stated that the brain structures responsible for musical anticipation are shaped by culture, personal listening history, and musical training (Vuust et al., 2005).

Armed with this knowledge, the following two hypotheses are formed:

1. It is hypothesized that non-diatonic chord progressions are perceived more negatively in terms of valence and higher in terms of arousal, and that the number of non-diatonic notes per progression correlates negatively with valence, and positively with arousal.
2. It is hypothesized that musicians are more consistent at rating different chord progressions in terms of valence and arousal compared with nonmusicians.

Methods

Participants

The study included 36 participants (13 females and 23 males) with a mean age of 39.1 (median = 27.5, SD = 18.3). Nearly all participants were from Finland and all were familiar with Western tonal music. Musical background was assessed before the actual experiment and those with more than three years of musical experience were recorded as musicians (N = 16), while the others were recorded as nonmusicians (N = 20).

Stimuli

The choice of stimuli for this study follows recommendations outlined by Eerola & Vuoskoski (2011) who point out that synthetic stimuli provide the researcher the ability to study and manipulate musical features. In contrast to real music, synthetic stimuli avoids the problem of participants being familiar with the sequences, with the trade-off of sounding artificial and lacking the intricate features of actual compositions (Eerola & Vuoskoski, 2011). For the purposes of this study, this trade-off was acceptable.

A tempo of 120 BPM was chosen, and the instrument used was a synthetic version of "cinematic strings" in the software application Logic Pro. The structure of each sequence was built as follows. First, a root note ("tonic") was randomly generated from one of the 12 semitones of the Western music system. Then either a major or minor triad was built on top, creating the tonic chord. The tonic chord was followed by another triad (hereafter referred to as the "shift chord") built on top of one of the 11 remaining semitones. The shift chord could also be either major or minor, with one additional rule governed by Western music theory. If the shift chord fulfilled the conditions of being minor and existing either on the seventh scale degree (11 semitones from the tonic) in a major key or the second scale degree (two semitones from the tonic) in a minor key, it was converted to a diminished chord by flattening the third note. For example, if C major was generated as the tonic chord and B minor as the shift chord, the shift chord would then be converted to B diminished. Each trial consisted of the tonic chord, shift chord, and tonic chord played in sequence. The length of one complete sequence was approximately six seconds.

Using the method described above, a total of 44 chord progressions can be created (four different pairs of major and minor, multiplied by 11 semitones). 12 out of these 44 progressions are diatonic, meaning that the shift chord belongs to the key signature dictated by the tonic chord. For the experiment, the total amount of chord progressions was reduced to 32 by removing 12 of the non-diatonic combinations, and thus ending up with all 12 diatonic combinations and 20 non-diatonic ones. The reasoning behind removing some of the combinations was the concern that participants might become desensitized to the nuances of the different chord progressions in case the experiment (which can be described as repetitive) continued for too long.

The 32 chosen progressions were presented to every participant in random order and in random key signatures. This means that all participants heard a "I – V – I" progression exactly once, but one participant might hear "C – G – C" while another participant might hear "Ab – Eb – Ab".

Procedure

Upon beginning of the experiment, the participants read through instructions displayed on-screen and had the opportunity to ask the researcher for clarification before continuing to the first trial. They were informed that their task was to rate different musical harmonies in emotional terms. Special care was also taken to ensure that two important distinctions in music research relevant to this study – the separation of perceived vs felt emotions, and ratings of valence vs preference (Schubert, 2007) – were considered. To specify, the participants were instructed to focus on perceived emotions and to disregard personal preference in their ratings.

The experiment consisted of 33 trials, the first of which was a practice trial intended to familiarize participants with the interface and keyboard controls of the experiment. At the beginning of each trial, participants heard a sequence of chords and were asked to rate their perception of the sound in terms of two dimensions of emotion, i.e. valence and arousal. Both dimensions were rated on continuous scales labelled "negative – positive" and "low energy – high energy" for valence and arousal respectively. An option to replay the sound once per trial was also available.

Results

The statistical analyses were conducted using the R statistical programming language and a confidence interval of .05 was used for all statistical tests.

Spearman's rank correlation coefficient was used to assess the relationship between the number of non-diatonic (outside) notes per chord progression and ratings of arousal and valence. While no significant correlation between outside notes and arousal was found (r_s = .03, p = .25, N =

36), a significant negative correlation between outside notes and valence was found, $r_s = -.20$, $p < .001$, $N = 36$. Thus, the first hypothesis was partially supported.

To assess whether musicians were more consistent in rating the chord progressions, intraclass correlation coefficient (ICC) estimates and their 95% confidence intervals were calculated using a two-way random-effects model, single-rating, and consistency. While the overall degree of reliability was low, it was found that musicians had higher scores for both ratings of arousal and valence. The ICC score for the arousal ratings of musicians was .0343 with a confidence interval from -.001 to .101, and for nonmusicians it was .022 with a confidence interval from -.005 to .075. The ICC score for the valence ratings of musicians was .41 with a confidence interval from .293 to .564, and for nonmusicians it was .204 with a confidence interval from .125 to .333. Therefore, the second hypothesis is supported by these findings.

Table 1. ICC scores

	Musicians	**Nonmusicians**
Arousal	.0343	.022
95% CI	*-.001 – .101*	*-.005 – .075*
Valence	.41	.204
95% CI	*.293 – .564*	*.125 – .333*

Figure 2.

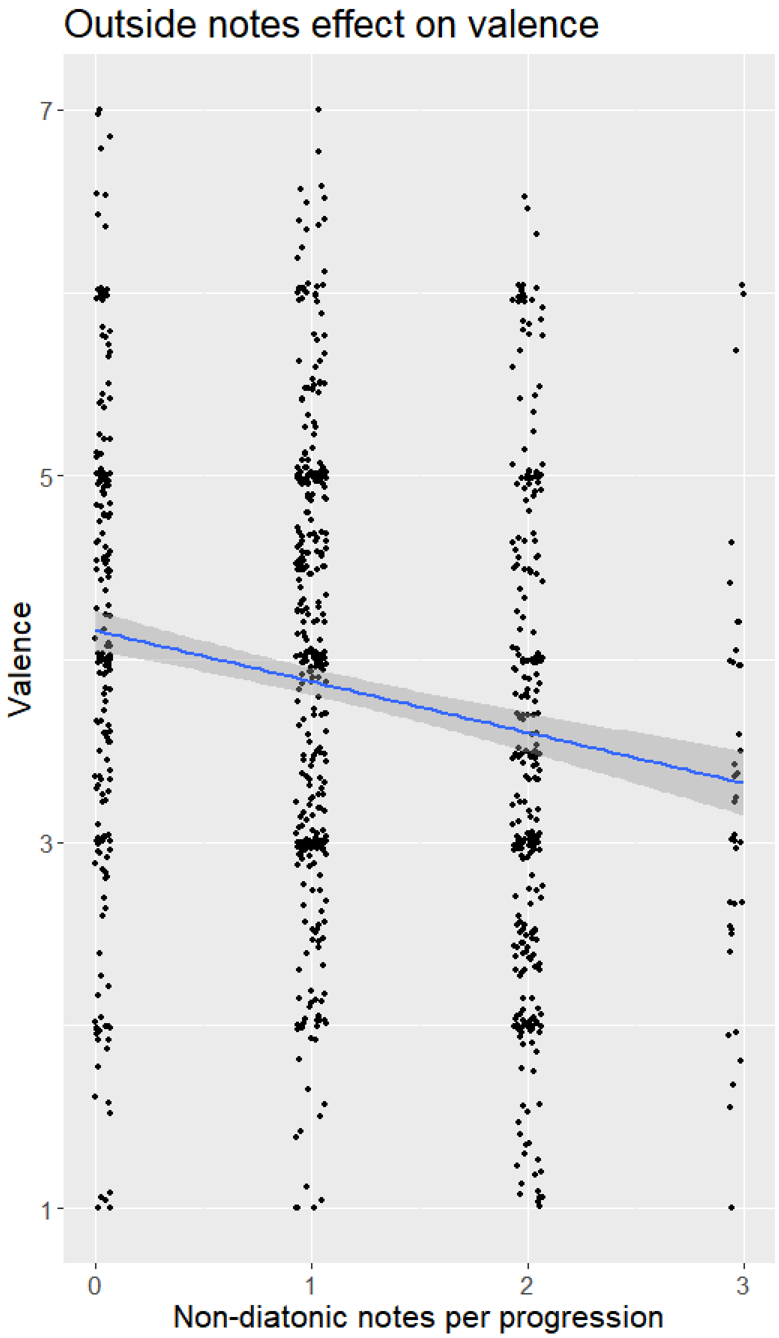
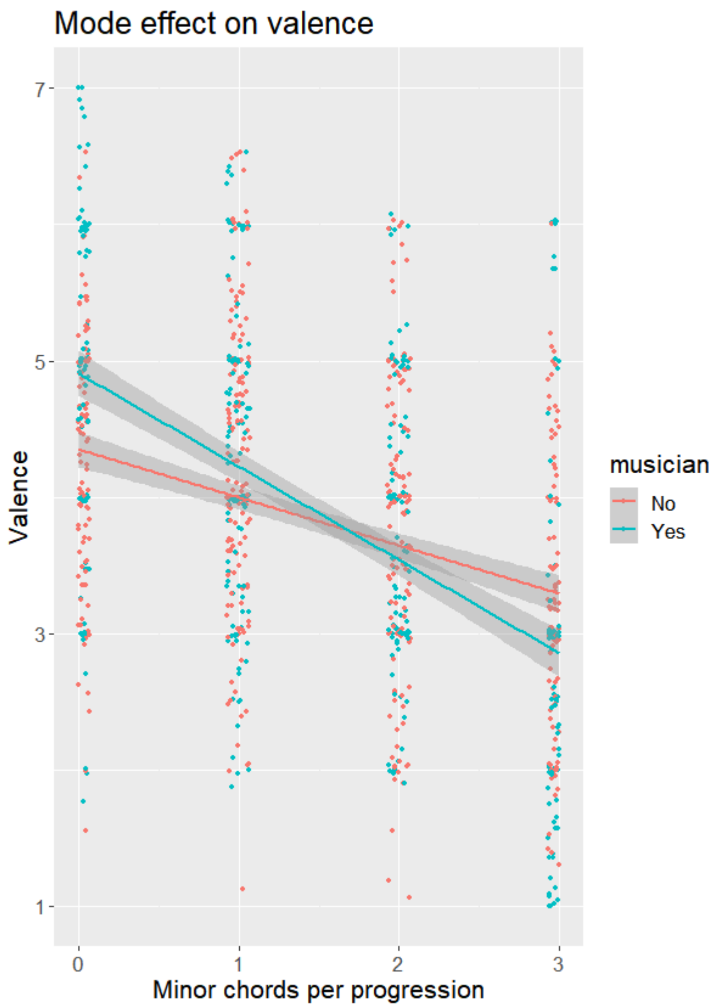

Outside notes effect on valence
Valence
7
5
3
1
0
1
2
3
Non-diatonic notes per progression

Figure 3.



Note: due to the structure of the progressions (e.g. i – III – i), a minor tonic chord added two to the minor count, while a minor shift chord added only one.

In addition, the relationship between mode and valence was assessed. Using Spearman's rank correlation coefficient, a significant negative correlation between minor chords and valence was found. This relationship was stronger for musicians (r_s = .55, p < .001, N = 16) than nonmusicians (r_s = .35, p < .001, N = 20).

Finally, the effect of pitch height on arousal and valence was measured using Spearman's rank correlation coefficient. Pitch height was found to correlate with both arousal (r_s = .31, p < .001, N = 36) and valence (r_s = .13, p < .001, N = 36).

Figure 4.

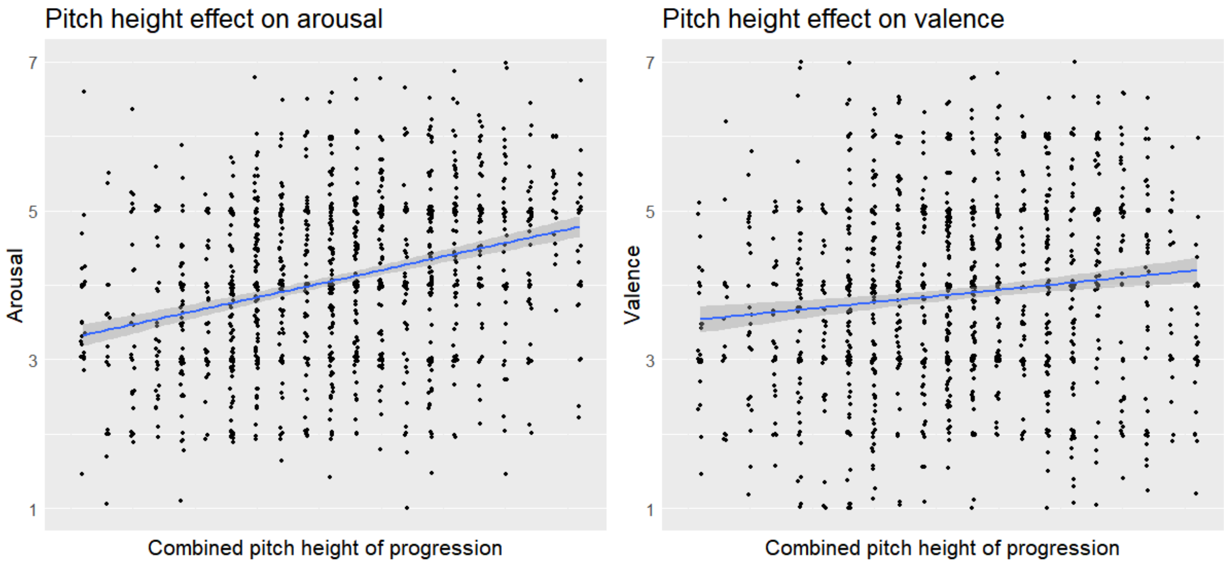
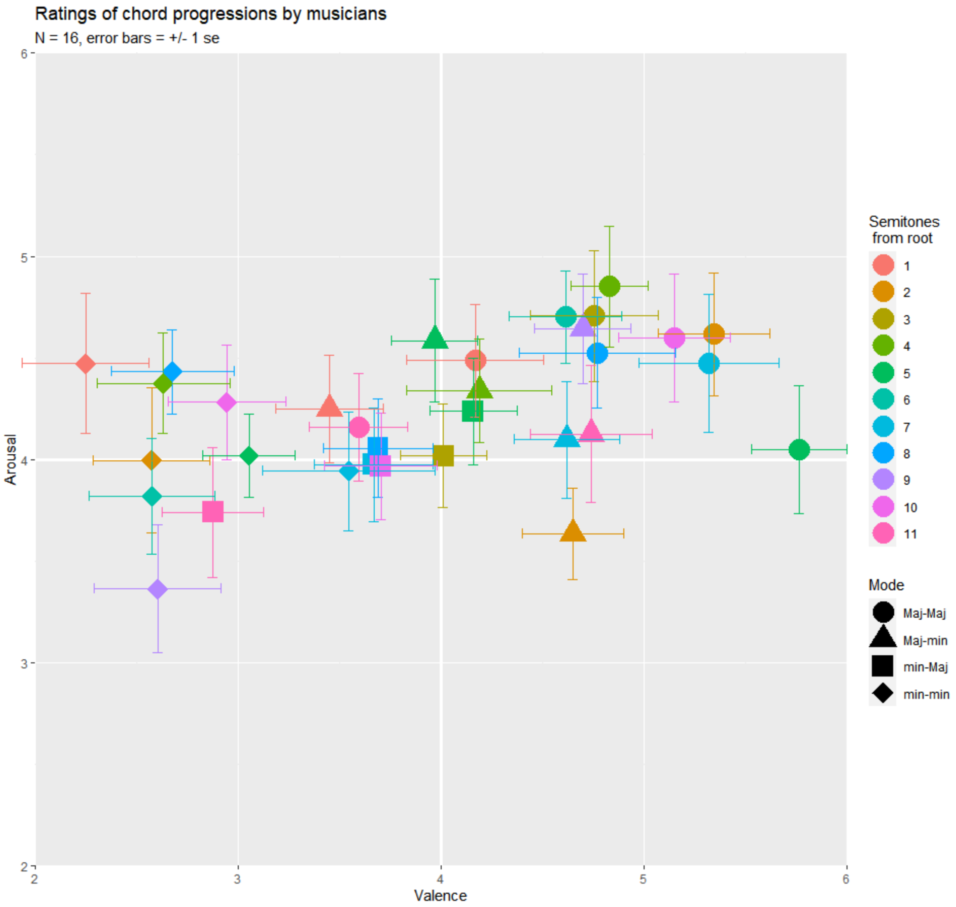
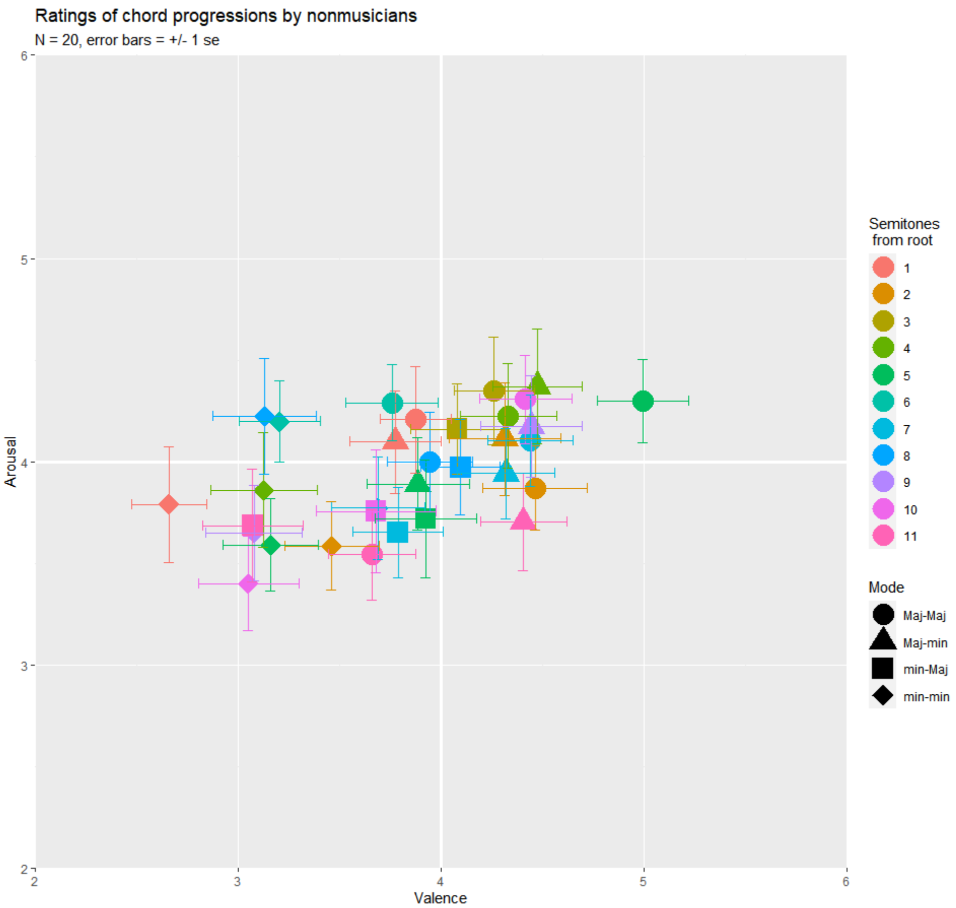

Pitch height effect on arousal
Arousal
7
5
3
1
Combined pitch height of progression
Pitch height effect on valence
Valence
7
5
3
1
Combined pitch height of progression

Figure 5.


Ratings of chord progressions by musicians
N = 16, error bars = +/- 1 se
Arousal
Valence
2
3
4
5
6
Semitones from root
1
2
3
4
5
6
7
8
9
10
11
Mode
Maj-Maj
Maj-min
min-Maj
min-min


Ratings of chord progressions by nonmusicians
N = 20, error bars = +/- 1 se
Arousal
Valence
2
3
4
5
6
Semitones from root
1
2
3
4
5
6
7
8
9
10
11
Mode
Maj-Maj
Maj-min
min-Maj
min-min

Discussion

Comparing the correlation coefficients of the different variables allows for some general conclusions to be made. It appears that a specific chord is perceived based in large part on its mode (major/minor), relative pitch, and relationship to the implied key signature (as measured by the number of non-diatonic notes). This holds true both for musicians and nonmusicians, which has some interesting implications. Nonmusicians are often consciously unaware of these differences and may find it difficult to say whether a chord belongs to the key signature or not – yet they systematically rate non-diatonic chords as more negative. The fact that diatonic chords (seen in the leftmost column of data points in figure 2) were rated higher in valence adds further support to the prediction effect outlined by Huron (2006).

Several things point to the fact that valence was a more intuitive way to rate the chord progressions compared to arousal. Figure 5 displays the mean ratings for each chord progression, first for musicians and then for nonmusicians. For both groups, it can be seen that the points are spread out more along the x-axis (representing valence) and clustered closer together along the y-axis (representing arousal). Additionally, the ICC scores (Table 1) show that reliability in ratings was higher for valence than arousal. It could be that arousal is affected more by the psychophysical dimensions of music, which would explain why the stimuli used in this study was unable to cause reliable differences in ratings of arousal. This is supported the fact that pitch height (Figure 4, left side), a psychophysical dimension, had the strongest effect on arousal (apart from valence, which is discussed in the limitations section).

The results also provided some insight into how previous experience affects perceived emotions. Musicians were more inclined to rate major chords as positive and minor chords as negative, which is illustrated by the steeper line for musicians in figure 3. In addition, musicians showed more consistency in their ratings, particularly for the valence-dimension.

Limitations

Arousal was found to be confounded with valence (r_s = .32, p<.001, N = 36), meaning that ratings tended to fall somewhere on the line between negative valence/low-energy and positive valence/high-energy. This may have been due to the configuration used in the study, which could have inadvertently included more chord progressions perceived to exist on the aforementioned line. Another explanation is that when participants were unsure about where to place the marker on the rating scale of one dimension, they roughly matched the rating on the other dimension. This is supported by the fact that the correlation between arousal and valence was stronger for non-musicians (r_s = .41, p<.001, N = 20) than musicians (r_s = .24, p<.001, N = 16), who may have been more confident in their ability to identify subtle differences in the harmonies. A frequent comment made by non-musicians after the experiment was “They all sound the same!”.

It is also worth mentioning that the musician group consisted of 14 males and two females. Although this study did not find significant effects of gender on perceived emotions, this imbalance could have impacted the results.

Finally, arousal was found to weakly correlate with the trial number (r_s = .07, p<0.001, N = 36), meaning that as participant got towards the end of the experiment, they tended to rate the stimuli slightly higher in terms of arousal.

Further research

While this study used three variations of triads – major, minor and diminished – further research could include other types of chords, such as seventh, suspended, and augmented. Melody could also be added as an independent variable, for instance by playing a melody in different key signatures or using different scales.

References

Balkwill, L.-L., & Thompson, W. F. (1999). A Cross-Cultural Investigation of the Perception of Emotion in Music: Psychophysical and Cultural Cues. *Music Perception: An Interdisciplinary Journal*, *17*(1), 43–64.

Bigand, E., Vieillard, S., Madurell, F., Marozeau, J., & Dacquet, A. (2005). Multidimensional scaling of emotional responses to music: The effect of musical expertise and of the duration of the excerpts. *Cognition & Emotion*, *19*(8), 1113–1139. https://doi.org/10.1080/02699930500204250

Collier, G. L. (2007). Beyond valence and activity in the emotional connotations of music. *Psychology of Music*, *35*(1), 110–131. https://doi.org/10.1177/0305735607068890

Eerola, T., & Vuoskoski, J. K. (2011). A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music*, *39*(1), 18–49. https://doi.org/10.1177/0305735610362821

Gomez, P., & Danuser, B. (2004). Affective and physiological responses to environmental noises and music. *International Journal of Psychophysiology*, *53*(2), 91–103. https://doi.org/10.1016/j.ijpsycho.2004.02.002

Gosselin, N. (2005). Impaired recognition of scary music following unilateral temporal lobe excision. *Brain*, *128*(3), 628–640. https://doi.org/10.1093/brain/awh420

Huron, D. (2006). Sweet Anticipation. Music and the Psychology of Expectation. *MIT Press.*

Ilie, G., & Thompson, W. F. (2006). A Comparison of Acoustic Cues in Music and Speech for Three Dimensions of Affect. *Music Perception*, *23*(4), 319–330. https://doi.org/10.1525/mp.2006.23.4.319

Khalfa, S., Delbe, C., Bigand, E., Reynaud, E., Chauvel, P., & Liéégeois-Chauvel, C. (2008). POSITIVE AND NEGATIVE MUSIC RECOGNITION REVEALS A SPECIALIZATION OF MESIO-TEMPORAL

STRUCTURES IN EPILEPTIC PATIENTS. *Music Perception*, *25*(4), 295–302. https://doi.org/10.1525/mp.2008.25.4.295

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*(3), 302–307. https://doi.org/10.1038/nn1197

Koelsch, S., and A. D. Friederici. (2003). Toward the neural basis of processing structure in music. Comparative results of different neurophysiological investigation methods. *Annals of the New York Academy of Sciences* 999: 15-28.

McDERMOTT, J., & Hauser, M. (2005). THE ORIGINS OF MUSIC: INNATENESS, UNIQUENESS, AND EVOLUTION. *Music Perception*, *23*(1), 29–59. https://doi.org/10.1525/mp.2005.23.1.29

Peretz, I. (2006). The nature of music from a biological perspective. *Cognition*, *100*(1), 1–32. https://doi.org/10.1016/j.cognition.2005.11.004

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, *17*(03). https://doi.org/10.1017/S0954579405050340

Schubert, E. (2007). The influence of emotion, locus of emotion and familiarity upon preference in music. *Psychology of Music*, *35*(3), 499–515. https://doi.org/10.1177/0305735607072657

Tellegen, A., Watson, D., & Clark, L. A. (1999). On the Dimensional and Hierarchical Structure of Affect. *Psychological Science*, *10*(4), 297–303. https://doi.org/10.1111/1467-9280.00157

Trehub, S. E. (2003). The developmental origins of musicality. *Nature Neuroscience*, *6*(7), 669–673. https://doi.org/10.1038/nn1084

Vieillard, S., Peretz, I., Gosselin, N., Khalfa, S., Gagnon, L., & Bouchard, B. (2008). Happy, sad, scary and peaceful musical excerpts for research on emotions. *Cognition & Emotion*, *22*(4), 720–752. https://doi.org/10.1080/02699930701503567

Vuust, P., & Kringelbach, M. L. (2010). The Pleasure of Making Sense of Music. *Interdisciplinary Science Reviews*, *35*(2), 166–182. https://doi.org/10.1179/030801810X12723585301192

Vuust, P., Pallesen, K. J., Bailey, C., van Zuijen, T. L., Gjedde, A., Roepstorff, A., & Østergaard, L. (2005). To musicians, the message is in the meter. *NeuroImage*, *24*(2), 560–564. https://doi.org/10.1016/j.neuroimage.2004.08.039

Witvliet, C. V. O., & Vrana, S. R. (2007). Play it again Sam: Repeated exposure to emotionally evocative music polarises liking and smiling responses, and influences other affective reports, facial EMG, and heart rate. *Cognition & Emotion*, *21*(1), 3–25. https://doi.org/10.1080/02699930601000672