

VALENCED PRIMING WITH ACQUIRED AFFECTIVE CONCEPTS IN MUSIC: AUTOMATIC REACTIONS TO COMMON TONAL CHORDS

IMRE LAHDELMA & TUOMAS EEROLA
Durham University, Durham, United Kingdom

THIS STUDY TESTED WHETHER CHORDS THAT DO not differ in acoustic roughness but that have distinct affective connotations are strong enough to prime negative and positive associations measurable with an affective priming method. We tested whether musically dissonant chords low in valence (diminished, augmented) but that contain little acoustic roughness are strong enough in terms of negative affective connotations to elicit an automatic congruence effect in an affective priming setting, comparable to the major-positive/minor-negative distinction found in past studies. Three out of 4 hypotheses were supported by the empirical data obtained from four distinct sub-experiments (approximately $N = 100$ each) where the diminished and augmented chords created strong priming effects. Conversely, the minor chord and the suspended fourth failed to generate priming effects. The results demonstrate how automatic responses to consonant/dissonant chords can be driven by acquired, cultural concepts rather than exclusively by acoustic features. The obtained results of automatic responses are notably in line with previous data gathered from self-report studies in terms of the stimuli's positive vs. negative valence. The results are discussed from the point of view of previous affective priming studies, cross-cultural research, as well as music historical observations.

Received: June 22, 2022, accepted October 19, 2023.

Key words: affective priming, chord, consonance, dissonance, familiarity

MUSICAL CONSONANCE AND DISSONANCE refers to the relative agreeableness/stability vs. disagreeableness/instability of simultaneous or successive pitch combinations (hereafter referred to as C/D and implying exclusively simultaneity). The exact acoustic and cultural components of C/D and their order of importance are a matter of ongoing debate, although during the last decade a consensus has

been emerging according to which C/D can be divided into three main constituents: the acoustic phenomena of *roughness* and *harmonicity*, and the cultural influence of *familiarity* (see e.g., Eerola & Lahdelma, 2021; Harrison & Pearce, 2020; Parncutt & Hair, 2011). Roughness denotes the sound quality that arises from the beating of frequency components (see e.g., Hutchinson & Knopoff, 1978), while harmonicity in turn indicates how closely a sonority's spectrum corresponds to a harmonic series (see e.g., Parncutt, 1989). Familiarity refers to the prevalence of sonorities in a given musical culture that affects how familiar listeners are with different chords and how this impacts perceived C/D (see Johnson-Laird et al., 2012; Lahdelma & Eerola, 2020); recent research has also drawn attention to the fact that harmonicity and cultural familiarity with Western music overlap significantly (Friedman et al., 2021; Lahdelma, Eerola, & Armitage, 2022).

The perception of C/D has been studied extensively with self-report methods (for an overview, see Harrison & Pearce, 2020), but automatic responses to it have been investigated considerably less. Linnavalli et al. (2020) found using an electroencephalographic mismatch negativity (MMN) method that sensory dissonance is discriminated at an early sensory level, irrespective of musical expertise. Affective priming studies by Sollberger et al. (2003), Steinbeis and Koelsch (2011), and Lahdelma, Armitage, and Eerola (2022) have found clear congruence effects for consonant/dissonant chords used as primes when categorizing positive/negative target words: when the chords and target words were congruent (positive-consonance or negative-dissonant), the affective categorization of the target words was significantly facilitated. All of these three affective priming studies found that musical expertise did not influence the responses. Recently Armitage et al. (2021) demonstrated that contrasts in roughness but not in harmonicity drive the automatic affective responses to consonant/dissonant intervals. Armitage et al. (2021) found that musically dissonant intervals (e.g., tritone and major seventh) that are typically rated unpleasant (see e.g., Bowling et al., 2018) do not create automatic negative congruence effects, but intervals within the critical bandwidth (see Plomp & Levelt,

1965) that contain high sensory dissonance (i.e., minor and major seconds) in turn do. The finding that roughness drives affective priming with intervals has implications for C/D research conducted on chords as well, as all of the above mentioned automatic reaction studies used only acoustically rough chord stimuli (including pitches from adjacent pitch classes causing interference between partials) to discriminate between consonance and dissonance. Importantly, unlike with intervals (cf. Armitage et al., 2021), chords have been shown to work even without differences in roughness as affective primes for valenced words, but so far only with regard to the major-positive/minor-negative affective dichotomy and not with regard to the question of C/D per se. This is an important incremental step forward to further our understanding of automatic reactions to C/D, as many musically dissonant chords such as the diminished and augmented chords are not acoustically rough.

With regard to the major/minor affective convention, Costa (2013) and Steinbeis and Koelsch (2011) used an affective priming method and found that major and minor chords facilitated the categorization of happy and sad words, respectively. Bakker and Martin (2015) presented valenced emotional facial stimuli simultaneously with major/minor chords and measured the facilitation of processing via event-related potential (ERP) amplitudes. When faces and chords were presented that contained congruent emotional information (major–happy or minor–sad), processing was facilitated. The authors concluded that major and minor chords indeed contain emotional connotations that can be processed as early as 200 milliseconds in naïve listeners. This major-positive/minor-negative mode distinction implies a culturally acquired (learnt) response rather than an acoustics-driven response due to a gradual acquisition (see Crowder et al., 1991; Dalla Bella et al., 2001; Gerardi & Gerken, 1995). Cross-cultural research points to a learnt response as well: Lahdelma et al. (2021) found a reversed self-reported affective response to major/minor chords (i.e., major-negative, minor-positive) among two tribes residing in remote Northwest Pakistan with minimal exposure to Western music, which they attribute to a reversed prevalence of major/minor sonorities in the tribes' music cultures compared to the West.

The purpose of the current study is to test whether chords that do not differ in roughness can have sufficiently distinct affective connotations (in addition to the major-positive/minor-negative convention) to prime negative and positive associations measurable with an affective priming method. We aim to investigate whether the distinction between C/D in chords is

indeed dependent only on roughness contrasts as it evidently is in the case of intervals (see Armitage et al., 2021)—in other words, we aim to test whether musically dissonant chords low in valence (i.e., diminished and augmented chords, see Lahdelma & Eerola, 2016a) but that contain little acoustic roughness are strong enough in terms of negative affective connotations to elicit an automatic congruence effect in an affective priming setting, comparable to the robustness of the major-positive/minor-negative distinction (see Bakker & Martin, 2015; Costa, 2013; Steinbeis & Koelsch, 2011) and the dichotomy between C/D stemming from large contrasts in roughness (see Lahdelma, Armitage, et al., 2022; Sollberger et al., 2003; Steinbeis & Koelsch, 2011) found in previous research. This is an important conceptual step to probe whether differences in automatic responses to consonant/dissonant chords are driven exclusively by roughness or whether these responses can also be acquired through cultural learning, as in the case of the above-mentioned major/minor affective distinction.

Acquired Affective Concepts of Harmony

To broaden the notion of culturally acquired affective priming beyond the major/minor distinction, we want to investigate automatic responses to some of the most foundational chords in Western music including all four *triads*. As Parncutt et al. (2019) point out, standard music theory texts, music dictionaries, and encyclopedias (see e.g., Piston & DeVoto, 1988; Randel, 1986) imply that Western tonality is based on four chords: the major, minor, diminished, and augmented triads. Despite not being acoustically rough or inharmonic (see Table 1) the diminished and augmented chords are considered musically dissonant due to the diminished fifth (tritone) interval contained in the diminished chord, and the augmented fifth interval contained in the augmented chord (Broyles, 1999; Persichetti, 1978). On account of their musical dissonance, the diminished and augmented chords communicate conventionally negative affect in the Western musical practice: Meyer (1956) proposes that these chords are often used to express intense emotion, apprehension, and anxiety. Indeed, Todd (1988) discusses how before the 19th century the augmented fifth was typically viewed as a passing dissonance, but through the works of Romantic era composer Franz Liszt the augmented chord emerged as an independent sonority and often denoted themes such as death, mourning, and grief. Similarly, Huckvale (1990) notes that in Richard Wagner's music the augmented chord frequently signifies distress. Tagg and Clarida

(2003) suggest that since its beginning film music has grounded its semantic practices on the European Classical and Romantic era—as film music is a major source for transmission of a culture’s musical conventions (Cohen, 2001), the negative affective connotations of the augmented and diminished chords (see also e.g., Kivy, 2002) form a clear music historical continuum. On an empirical note, Lahdelma and Eerola (2016a) have demonstrated that both diminished and augmented chords are perceived notably low in valence among Western listeners, while cross-cultural research has found an indifference with regard to preference of the major and the augmented chords (McDermott et al., 2016) or even a preference for the augmented chord over the major chord (Lahdelma et al., 2021) in non-Western populations, implying a culturally acquired aesthetic response.

The negative affective connotations of these chords may also be related to their low frequency of occurrence in actual music as per the mere exposure effect that postulates that familiarity yields positive valence (Zajonc, 2001). Previous corpus studies demonstrate the contrast in the prevalence of chord types across many musical styles. Rohrmeier and Cross (2008) found that the most prevalent chord type in Bach chorales in both major and minor modes is the major chord (60.8% in major modes, 44.9% in minor modes), followed by the minor chord (17.1% in major modes, 33.8% in minor modes), while the diminished chord accounts for only 2.3% in major modes and 3.3% in minor modes. Broze and Shanahan (2013) in turn found that in a jazz corpus of pieces from 1924 to 1968 major chords have a frequency of 22.1%, while diminished and augmented chords constitute only 2.2% and 0.1% of the corpus, respectively. With regard to popular music, De Clercq and Temperley (2011) created a rock harmony corpus from 100 songs published between the 1950s and the 1990s and found that 75.8% of chords in the corpus are major, 23.4% are minor, 0.7% are diminished, and 0.1% are augmented. Burgoyne (2012) in turn constructed the *Billboard Data Set* (a large corpus of music sampled from the US charts published between 1958 and 1991) and found that major chords have a frequency of 51.6%, minor chords 13.0%, and diminished chords 0.2%. The augmented chord’s percentage was not elaborated due to its rarity. However, according to previous research some positively valenced and preferred chords such as the major ninth and the major pentatonic (alternatively named “sixth/ninth” or “maj6add9”) chords (Lahdelma & Eerola, 2016b) are also exceedingly rare in frequency (0.4% and 0.2% respectively in the *Billboard Data Set*) and hence this rationale may be of limited value—we

surmise that negative valence in chords is driven by learnt associations rather than by a linear correlation with frequency of occurrence (cf. the mere exposure effect, Zajonc, 2001).

Hence, we also want to test whether negative affect in musically dissonant chords is driven more by established affective conventions than by low familiarity per se. For this end we include the suspended fourth (“sus4”) as a control chord that is low in familiarity (0.9% in the *Billboard Data Set*). It is typically seen as an alteration of a major or minor chord as opposed to being a chord in its own right, and within the European classical tradition it conventionally resolves onto a major or minor chord (Parncutt et al., 2019). Despite its low familiarity the suspended fourth has been rated as more pleasant than the augmented and diminished chords in previous empirical experiments (Arthurs et al., 2018; Bowling et al., 2018).

Hypotheses

We have four hypotheses and treat each of these as a separate sub-study that will utilise identical methodology. The specific hypotheses are as follows:

1. **Major - Minor:** Affectively congruous target words (major-positive/minor-negative) will be categorized faster than affectively incongruous (major-negative/minor-positive) target words in line with previous affective priming experiments (Costa, 2013; Steinbeis & Koelsch, 2011).
2. **Major - Augmented:** Congruous target words (major-positive/augmented-negative) will be categorized faster than incongruous (major-negative/augmented-positive) target words due to the augmented chord’s low valence, as demonstrated with self-reports (Arthurs et al., 2018; Lahdelma & Eerola, 2016a).
3. **Major - Diminished:** Congruous target words (major-positive/diminished-negative) will be categorized faster than incongruous (major-negative/diminished-positive) target words due to the diminished chord’s low valence, as demonstrated with self-reports (Arthurs et al., 2018; Lahdelma & Eerola, 2016a). As the diminished chord is the least harmonic of the current stimuli (see Table 1), if the priming results for this pair of chords are substantially larger than for the more harmonic dissonant chords (aug, sus4), the implied contribution of harmonicity will require a separate empirical design with controlled levels of roughness and harmonicity.

4. **Major - Suspended Fourth:** No congruence effects; despite its relatively low familiarity (see Table 1) the suspended fourth has been rated as more pleasant than the augmented and diminished chords in previous self-report experiments (Arthurs et al., 2018; Bowling et al., 2018) rendering it an ineffective prime for negative target words.

We treat these four distinct hypotheses as separate sub-studies that will utilise identical methodology.

Method

The methods and the analyses were preregistered and can be found at <https://osf.io/x8yvk>. The current experiment uses an affective priming method. This method is common in both social and cognitive psychology as an indirect measure of attitudes (for an overview see e.g., Herring et al., 2013). The affective priming paradigm consists of two stimuli (the prime and the target) presented in quick succession. The extent to which the first (prime) stimulus influences responses to the second (target) stimulus is indexed by a reaction time or accuracy rate. Target stimuli are typically evaluated more quickly when preceded by a prime of the same affective category (congruent) compared to when preceded by one of the opposite category (incongruent). This is surmised to be due to the prime stimuli activating responses based on valence (De Houwer et al., 2009). Affective priming paradigms have been consistently found to be a robust measure of attitudes to prime stimuli (see e.g., Fazio, 2001) and it allows to study affective responses via automatic, objective means. Research on C/D is prone to confusion emerging from semantic labels and priming offers a possibility to bypass such issues. Affective priming can tell us whether C/D discrimination on an automatic level is exclusively due to roughness contrasts or not. Previous research has successfully demonstrated the value of automatic

affective responses in differentiating these responses to chords, so far without disentangling the underlying mechanisms (see Costa, 2013; Lahdelma, Armitage, & Eerola, 2022; Sollberger et al., 2003; Steinbeis & Koelsch, 2011).

STIMULI

Chords

The experiment stimuli consisted of all four triads (major, minor, diminished, augmented) and the suspended fourth chord as a control chord (see Table 1). The roots of all chords were doubled in the octave to keep the outer voices (bass and soprano) still (see Lahdelma, Armitage, & Eerola, 2022; Sollberger et al., 2003; Steinbeis & Koelsch, 2011 for a successful implementation in a reaction time setting); previous literature points out that there is special perceptual salience for outer voices (see e.g., Bigand et al., 1996; Huron, 2001), which would pose a possible confound in alternating chord types unless kept constant.

To briefly describe the known musical and acoustic properties of the chord stimuli, Table 1 summarizes acoustic analyses and empirical pleasantness ratings for the chords. All chords are low in roughness (measured with Hutchinson & Knopoff, 1978), and high or moderately high in harmonicity (measured with Stolzenburg, 2015). Apart from the major and minor chords, the chords are infrequently used in actual music (< 1% in a popular music corpus, see Burgoyne, 2012, p. 163). All chords have previously been rated relatively pleasant except for the diminished and augmented chords (see Bowling et al., 2018). To contextualise the results of the acoustic models and the pleasantness ratings for the stimuli, we have offered the quantiles (shown in brackets) of these variables calculated from all possible 220 four-pitch combinations that can be formed within an octave (see the Durham Chord Dataset, Eerola & Lahdelma, 2021).

The chord stimuli were generated with *Ableton Live 9* (a music sequencer software) using the *Synthogy Ivory*

TABLE 1. *Chord Properties*

Name	Pitch-class	Roughness	Harmonicity	Frequency	Pleasantness
Major	{0,4,7,12}	0.19 (<25%)	-2.00 (>75%)	51.6%	3.83 (>75%)
Minor	{0,3,7,12}	0.19 (<25%)	-3.32 (>75%)	13.0%	3.47 (>75%)
Augm	{0,4,8,12}	0.27 (<25%)	-3.42 (>75%)	<0.2%	2.37 (>50%<75%)
Dim	{0,3,6,12}	0.21 (<25%)	-4.32 (>50<75%)	0.2%	2.87 (>75%)
Sus4	{0,5,7,12}	0.22 (<25%)	-3.02 (>75%)	0.9%	3.50 (>75%)

Note: Quantiles (in brackets) calculated from all possible 220 four-pitch combinations that can be formed within an octave (see the Durham Chord Dataset, Eerola & Lahdelma, 2021) for roughness (Hutchinson & Knopoff, 1978), harmonicity (Stolzenburg, 2015), chord frequency (Burgoyne, 2012, p. 163), and the pleasantness ratings of all 220 four-pitch combinations obtained from Bowling et al. (2018).

Grand Pianos II plug-in with *Steinway D Concert Grand* as the applied sound font, with a fixed velocity of 65. All chords were played in equal temperament as per previous affective priming studies using chords as stimuli (see e.g., Costa, 2013; Lahdelma, Armitage, & Eerola, 2022; Sollberger et al., 2003). We chose the register of the stimuli to be optimal by eliminating additional roughness and sharpness arising from low and high registers (see Eerola & Lahdelma, 2022): all chords were centered around A_3 and with transposed versions (± 5 semitones, all transpositions being equally likely to occur within this range) of each to eliminate any sense of tonal progressions. The duration of the sounds was 800 ms, which included a 10 ms fadeout to prevent any artifacts. The loudness was normalized to -10 dB peak using the ITU-R BS.1770-4 protocol (Steinmetz & Reiss, 2021). The roughness and harmonicity of the chords, as measured with models by Hutchinson and Knopoff (1978) and Stolzenburg (2015), respectively, is shown in Table 1. The calculation of these measures was implemented through the *incon* library by Harrison and Pearce (2020).

Words

All words were obtained from Warriner et al. (2013) containing ratings for valence, arousal and dominance. We chose words that are similar in lexical frequency and length (5 ± 1 letters). The positive words—with their valence ratings shown in brackets—were *sweet* (7.77), *lively* (7.12), *gentle* (7.42), *cuddle* (7.60), *excite* (7.79), *kiss* (7.78), *comfy* (7.25), and *relax* (7.82); the negative words were *rabid* (2.95), *hijack* (1.84), *death* (1.89), *vomit* (1.98), *arrest* (2.33), *fatal* (2.00), *dismal* (2.60), *morgue* (1.79). The negative and positive words do not differ in terms of arousal (Negative $M = 5.44$, $SD = 1.44$, Positive $M = 4.39$, $SD = 1.64$, $t = 1.35$, $p = .199$) nor lexical frequency (Negative $M = 37.6$, $SD = 74.9$, Positive $M = 48.8$, $SD = 61.4$, $t = -0.325$, $p = .75$) using word frequency per million words (Van Heuven et al., 2014).

PROCEDURE

We collected the data using *PsyToolkit*, a web-based service designed for reaction-time experiments (Stoet, 2010, 2017). The participants were required to use headphones to take the experiment; they had to pass (score at least 5 out of 6 items correct) a headphone check proposed by Woods et al. (2017) to proceed to the experiment. The musical sophistication of the participants was collected with the 1-item self-report measure titled *Ollen Musical Sophistication Index* (OMSI) by Ollen (2006) that asks the participant to identify with one of the six choices presented (nonmusician, music-

loving nonmusician, amateur musician, serious amateur musician, semiprofessional musician, or professional musician).

The experiment was a standard word classification task with affective priming. Each item consisted of the prime (chord) presented simultaneously with a fixation cross for the duration of 200 ms. At 200 ms, the fixation cross disappeared to be replaced with the target word. Participants were instructed to press the “z” key if the target word is negative and the “m” key if it is positive on their keyboard. The target word was visible onscreen for the duration of 1500 ms; key presses longer than 2000 ms after the onset of the target word were classed as timeouts. Participants first completed a six-item familiarisation block, which was followed by the experimental block of 64 items. During the practice block, participants were informed whether or not their response was correct immediately after each item. No indication of accuracy was revealed during the experimental block.

SAMPLE SIZE

Previous studies have used between 20 and 76 participants in affective priming experiments. Armitage et al. (2021) had an average of 40 participants in each sub-study (10 experiments in total), Sollberger et al. (2003) had 43 (in Experiment 1) and 76 (in Experiment 2), Lahdelma, Armitage, and Eerola (2022) 40, Steinbeis and Koelsch (2011) 20 (in Experiment 1), 20 (in Experiment 2) and 33 (in Experiment 3), and Costa (2013) 70 (in Experiment 1), 41 (in Experiment 2), and 41 (in Experiment 3). In the current study, we wanted to keep the experiment length compact and compensated for the number of observations with increasing the sample size compared to the related studies. For our design (4 conditions \times 16 words = 64 items) for one chord pair (e.g., Major-Minor), we collected data from 100 participants as this created a comparable number of observations per cell to other studies. In addition, we calculated the power analysis and sample size estimation that supports the proposed N (reported in a later section). We recruited participants from Prolific.ac.uk with two recruitment criteria: 1) English as a native language to ensure a clear understanding of the instructions and the target words (see Tenderini et al., 2022), and 2) right-handedness of the participants (see Hardie & Wright, 2014).

Responses which were too rapid (< 200 ms), failed to respond, or responded too slowly (> 1500 ms) or incorrectly (Brysbaert & Stevens, 2018) were discarded. We recruited initially 100 participants per sub-study to obtain the planned sample size (100, see the following

TABLE 2. Sample Characteristics

Sub-experiment	N	Gender (F/M/O)	Age M (SD)	OMSI (NMI/MI/SMI)	Correct
Maj-Min	100	49%/46%/5%	25.3 (3.3)	82%/18%/0%	95.9%
Maj-Aug	102	48%/47%/5%	24.6 (3.2)	72%/25%/3%	95.2%
Maj-Dim	99	42%/57%/1%	25.3 (3.3)	76%/24%/0%	99.0%
Maj-Sus4	101	50%/49%/1%	25.7 (2.7)	75%/22%/3%	98.1%

Note: Description of sample size, gender, age, musical expertise (OMSI 1-item question classified into No Musical Identity (NMI), Musical Identity (MI), and Strong Musical Identity (SMI)), and the proportion of correct responses in the affective priming task for each sub-experiment.

section). After this recruitment, we checked the quality (missed answers, the speed of responses) to determine how many of the participants did not fulfill the set criteria. When we did not have the target sample size after eliminating those that did not fulfill the criteria, we recruited the missing number of participants by requesting 3–5 more participants to a sub-experiment, resulting in near or exact sample sizes of 100 (see Table 2 for the exact numbers). Overall, we collected data for four sub-studies: each involved the same paradigm, similar procedures, similar amount of participants, but a different chord pairing.

SAMPLE DESCRIPTION

The participants in all four sub-experiments ($N = 402$) were mainly those without musical identity (76%) as measured by a single item OMSI measure that was split into three categories as suggested by Zhang and Schubert (2019)—No Musical Identity (NMI), Musical Identity (MI), and Strong Musical Identity (SMI), and reported in Table 2. Nearly an equal proportion of women (47.3%) and men (49.5%) and a minor proportion (3.2%) of those who did not either disclose gender information or identified with non-binary classification participated in the sub-experiments. The participant ages ranged from 18 to 30 with a mean of 25.2 years ($SD = 3.1$). These basic demographics were similar across the sub-experiments (evidenced by the lack of differences for gender ($\chi^2 = 12.5$, $p = .67$), age ($F[3, 398] = 1.95$, $p = .12$), and musical expertise, $\chi^2 = 3.82$, $p = .30$). Table 2 shows the breakdown of the sample characteristics for each sub-experiment as well as the proportion of correct responses given in the tasks after the elimination of participants that failed more than 80% of the trials (altogether 9 participants).

Analysis Strategy

In each sub-study, our study design was a within-participant experiment, where we manipulated *Valence* by priming it with *Words* and *Chords*, which formed

either congruent (positive word - consonant chord or negative word - dissonant chord) or incongruent pairs (positive word - dissonant chord, or negative word - consonant chord). For the analysis, this is a within subjects design with one factor: *Congruence* with 2 levels (Congruent and Incongruent, which come from collapsing the four conditions of Words and Chords). All participants made decisions on 64 stimuli. We considered participants as a Random factor. The stimulus presentation and the pitch range were randomized (random transposition of ± 5 semitones from A_3), where each participant were presented with all pairings of Chord and Words twice in a random order. The randomization was built into the experimental interface.

```
model <- brms::brm(RT ~ congruence +
  (1 + congruence | participant) +
  (1 | transposition) + (1 | word),
  data = substudy1,
  family = 'shifted_lognormal',
  iter = 4000, chains = 4, thin = 1,
  prior = prior(student_t(1, 0.00, 0.71),
  coef = congruence),
  )
```

The results were analyzed in R (R Core Team, 2022) and we used the brms package (Bürkner, 2021) to estimate the parameters of the model. We specified the distribution of reaction times with a log-related function (shifted_lognormal) appropriate for RT responses and specified an informative prior for congruence (*Student's t* with 1 degrees of freedom and $\mu = 0.00$, and $\sigma = 0.71$) based on previous data (Lahdelma, Armitage, & Eerola, 2022) where these parameters offered clear evidence of congruence impacting reaction times (median $\theta = 0.0155$, where lower 95% CI = 0.007, upper 95% CI = 0.0253). We specified random slopes for congruence for each participant. We also incorporated random intercepts for transpositions and words (Barr et al., 2013). Following Kruschke (2014), the sample size was

determined by calculating the 80% probability that a study design obtains an effect size of 0.35 with the Bayes factor over 6. This estimation, executed with the BFDA package (Schönbrodt & Wagenmakers, 2018), with modest effect size (Cohen's $D = 0.35$) and within-subject analysis suggested that 80% power is to be achieved at $N = 100$, which gives 80.1% for strong evidence for H_1 ($BF > 6$), and the rest 19.7% indicate weak support for the hypothesis ($0.1667 < BF < 6$), with 0.2% evidence for H_0 ($BF < 1/3$). Traditional power analysis and sample size estimation yields $N = 110$ with a power of 80 using R library Superpower (Lakens & Caldwell, 2021). When the effect of *Congruence* was assessed with Bayesian evidence, we also described the relation between Chord and Word congruence to demonstrate any asymmetries. Scripts and data are available at Github https://github.com/tuomaseerola/cultural_priming.

In terms of demographic background questions, we collected *age*, *gender*, and *musical sophistication* using the Ollen Musical Sophistication Index's 1-item measure (for a rationale, see Zhang & Schubert, 2019) from all participants.

The initial data screening involved discarding participants who (a) failed the headphone check (approximately 6–8% of the sample in each sub-experiments), (b) failed to respond to all items (failed to complete the task, less than 2% of the participants), and (c) responded too fast or too slow ($< 200\text{ms}$ or $> 1500\text{ms}$) which were eliminated on trial-by-trial (1.9%–4.8% of the trials, see Table 2), and (d) responded incorrectly (participants were eliminated if incorrect for more than 80% of the responses, a total of 9 participants). The experiment was approved by the ethics committee of the Department of Music at Durham University and was conducted in accordance with its guidelines and regulations (MUS-2022-03-21T14_47_40-lqbn73).

Results

Our main interest was to assess whether congruency (congruent or incongruent valenced targets and primes) led to differentiated reaction times (population-level effects) while allowing individual slopes for congruence for each participant and random intercepts for transposition and words (group-level effects). As the Bayesian paradigm offers insight into the evidence based on observed data and a prior distribution, and estimated posterior probability, we report the *credible intervals* for the specific hypothesis of our design (Kruschke, 2018). Out of the group-level effects, Random Transposition

TABLE 3. Descriptives Across Conditions

Sub-experiment	Target	Prime	
		Neg. Chord	Pos. Chord
Min	Neg. Word	550 (546–554)	555 (551–559)
	Pos. Word	542 (538–546)	540 (536–545)
Augm	Neg. Word	557 (553–561)	565 (561–569)
	Pos. Word	550 (546–555)	550 (545–554)
Dim	Neg. Word	545 (540–549)	552 (548–556)
	Pos. Word	543 (539–548)	538 (534–543)
Sus	Neg. Word	545 (540–549)	547 (543–539)
	Pos. Word	535 (530–539)	533 (528–537)

Note: Mean and $CI_{95\%}$ in brackets.

and Participant produced coefficients that randomly varied around 0, and Words showed a pattern consistent with their valence across the sub-experiments (full breakdown available in the online materials).

Starting with descriptive summaries, Table 3 shows the mean reaction times and 95% confidence intervals across prime (chords) and target (words) for each sub-experiment. Overall, we see a large effect of negative words (targets) being rated more slowly than positive words but no clear difference between primes (two chords). The strongest suggestion of interaction between target and prime seems to take place in the sub-experiment with the Major-Diminished chord pairing. In this case we see a typical affective priming effect where the incongruent pairing of the prime and target results in a slower reaction time. In other words, the major chord followed by negative words compared to the diminished chord followed by negative words has a difference of +7 ms ($552_{PosPrime} - 545_{NegTarget} = 7$), and the diminished chord followed by positive words compared to the major chord followed by positive words yields a delayed reaction time of +5 ms ($543_{NegPrime} - 538_{PosTarget} = 5$), see Table 3. In general, the differentiated speed of reaction is more apparent with negative target words than in positive ones (Gao et al., 2020; Scherer & Larsen, 2011). However, it must be remembered that summarizing the means across conditions does not show fully the extent to which individual responses are affected by the affective priming paradigm, and in the actual analyses we also incorporate several group-level effects into the model (individuals, transpositions, and the words themselves). Also the distribution of the reaction times is not particularly well described with the means as it is non-normally distributed, which is also taken into consideration in the analysis by modelling the responses with a log-related function.

In these sub-experiments, we are specifically looking for the interaction between target and prime (the effect

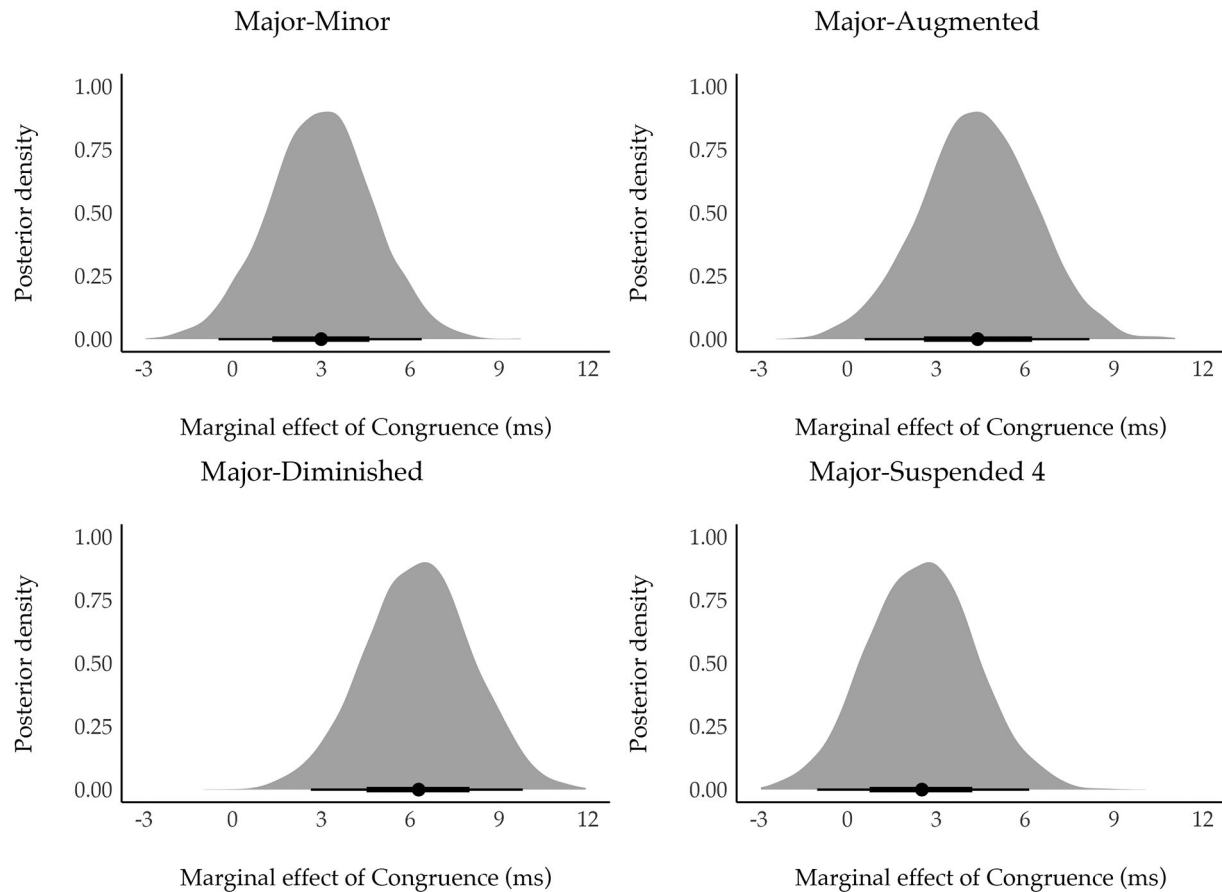


FIGURE 1. Marginal effects of Congruence for the four sub-experiments. The grey area marks the 95% credibility interval.

of congruence), which is best visualized by the the marginal effects of congruence (Figure 1), the posterior distribution of congruence (Figure 2), and the directed hypothesis concerning the congruence. Figure 1 pools together visually the size of the effect of congruence for each sub-experiment, which shows clear marginal effects of congruence across sub-experiments, the median values ranging from 3.0 ms ($CI_{95\%}$ -0.4–6.4 ms) for the minor chord, 4.4 ms ($CI_{95\%}$ 0.5–8.1 ms) for the augmented, to 6.3 ms ($CI_{95\%}$ 2.8–9.9 ms) for the diminished, and to 2.5 ms ($CI_{95\%}$ -1.1–6.0 ms) for the suspended fourth.

Table 4 reports the summary of the population-level effects for each sub-experiment where we test whether the effect of congruence is greater than zero. The estimate provides the median value of the posterior (θ) distribution, *Estimate Error* is the standard deviation of θ and we use a conservative 95% credible interval (*CI Lower* and *CI Upper*) as the critical range of the posterior values. If the credible interval of the

posterior distribution does not contain zero, the analysis suggests that the effect of congruence has positive evidence for the hypothesis. The weight of the evidence can be quantified in different ways; one suitable for our directed hypotheses is the *probability of direction* (pd), which is the proportion of the posterior distribution that is of the median's sign (Makowski et al., 2019). It varies between 50% and 100% and the values have similarity to frequentist p values through transformation of $1-pd$ (one-sided p value) and $2(1-pd)$ where pd values under 95% are $p > .10$, $pd > 95\%$ are $p < .10$, $pd > 97\%$ are $p < .06$, $pd > 99\%$ are $p < .02$, and $pd > 99.9\%$ are $p < .002$ (*Bayesian Reporting Guidelines*, n.d.).

A breakdown of the Bayesian analysis of the main hypothesis for each sub-experiment in Table 4 confirms what the visualization of the marginal effects (Figure 1) already indicated. Starting with the sub-experiment with a minor chord paired with a major chord, the

TABLE 4. Summary of Bayesian Analysis and Directed Hypothesis Testing

Comparison	Hypothesis	Estimate	Est. Error	CI_{Lower}	CI_{Upper}	ER	pd (%)
Maj-Min	Congr. > 0	0.0035	0.0020	0.0001	0.0069	20.7	95.4*
Maj-Aug	Congr. > 0	0.0047	0.0021	0.0013	0.0081	79.8	98.8*
Maj-Dim	Congr. > 0	0.0070	0.0021	0.0036	0.0105	3999	100*
Maj-Sus4	Congr. > 0	0.0026	0.0019	-0.0005	0.0057	11.1	91.8

Note: * refers to the expected value under the hypothesis falling outside the credibility interval ($CI_{95\%}$), estimate is the median value and error is the standard deviation of the posterior distribution, CI lower and CI upper refer to 95% credible intervals. Evidence ratio (ER) denotes the posterior probability under the hypothesis against its alternative, and probability of direction (pd (%)) is the certainty associated with the positive direction of the effect.

credible interval of the posterior distribution includes 0 (median θ of 0.0035, $CI_{95\%}$ of 0.0001–0.0069) so there is only weak to non-existing support for the effect of priming even though the majority of the θ distribution lies above zero. The positive direction of the effect (pd = 95.4%) suggests a marginal or weak effect (see Table 4 and also Figure 2 for the posterior distributions). The interpretation is that the minor chord is not sufficiently strong to generate measurable affective priming effect when paired with the major chord and with negative and positive words, respectively. For the augmented chord, the results show strong positive evidence for congruence (median θ of 0.0047, $CI_{95\%}$ of 0.0013–0.0081) where the credible interval does not contain zero and the evidence ratio is above 30 (79.8) and the positive direction of effect shows strong certainty with the associated effect (pd = 98.8%). Moving on to the diminished chord, the analysis provides strong evidence for the effect of congruence (median θ of 0.0070, $CI_{95\%}$ of 0.0036–0.0105) where the credible interval does not contain zero. The evidence ratio is very large (3999) and the positive direction of the effect shows a resounding confidence of the interpretation (pd = 100%). Finally, the suspended fourth chord shows lack of persuasive evidence of congruence as the credible interval for posterior probability contains 0 (median θ = 0.0026, $CI_{95\%}$ of -0.0005–0.0057) and the direction of effect is uncertain (pd = 91.8%).

Figure 2 displays the distribution of the posterior probabilities (θ) with the 95% credible interval and 0 highlighted for each sub-experiment. The successful sub-experiments contain differences between the congruence that are in the magnitude of 4.4 ms ($CI_{95\%}$ 0.5–8.1) for the augmented chord and 6.3 ms ($CI_{95\%}$ 2.8–9.9) for the diminished chord. In terms of congruence effects, these are both in the lower threshold of successful affective priming effects with words and pictures (Spruyt et al., 2002) and in the same magnitude as several past affective priming studies with chords and intervals (Armitage et al., 2021; Lahdelma, Armitage, & Eerola, 2022).

Connecting the observations to our hypotheses, the *Major – Minor* pairing with positively and negatively valenced words received only weak support for our hypothesis that these chords have clear negative and positive connotations. The credible interval contains 0 and the interpretation of the pd of 95.4% in frequentist statistics would put the p value between .05 and .10. The *Major – Augmented* pairing produced a clear priming effect, as hypothesized. The *Major – Diminished* pairing produced a strong priming effect consistent with our hypothesis. Finally, the *Major – Suspended Fourth* pairing was not predicted to generate priming effects and the empirical evidence supported this hypothesis (95% credible interval contains 0 and the direct hypothesis produces probability of direction of 91.8%, which is interpreted as not giving support for the hypothesis). In the online supporting materials, we also provide auxiliary analyses with linear mixed models that have the same population and group-level factors as the present analysis but follows the frequentist paradigm (lme4, Bates et al., 2015) and offers an identical interpretation of the data for each sub-experiment.

Finally, how large are the effects of the two successful primer chords, namely the augmented and diminished chords? This is an important question as it addresses the potential role of harmonicity in explaining the results; the diminished chord is the only chord in the current stimuli that did not lie in the top 25% quantile of harmonicity values and while it cannot be considered to be *inharmonic* (it lies in the 50%-75% quantile distribution of harmonicity, see Table 1), the diminished chord did obtain the largest marginal effects of congruence. To assess the difference when compared to the augmented chord, we use the posterior distributions of the sub-experiments with the diminished and augmented chords to estimate the size of this difference. We first calculate the difference between the two posterior distributions, which returns a normal distribution. From this we can tally the positive side of the distribution, which indicates that there is a posterior probability of 79.4% that the congruence for the diminished chord is

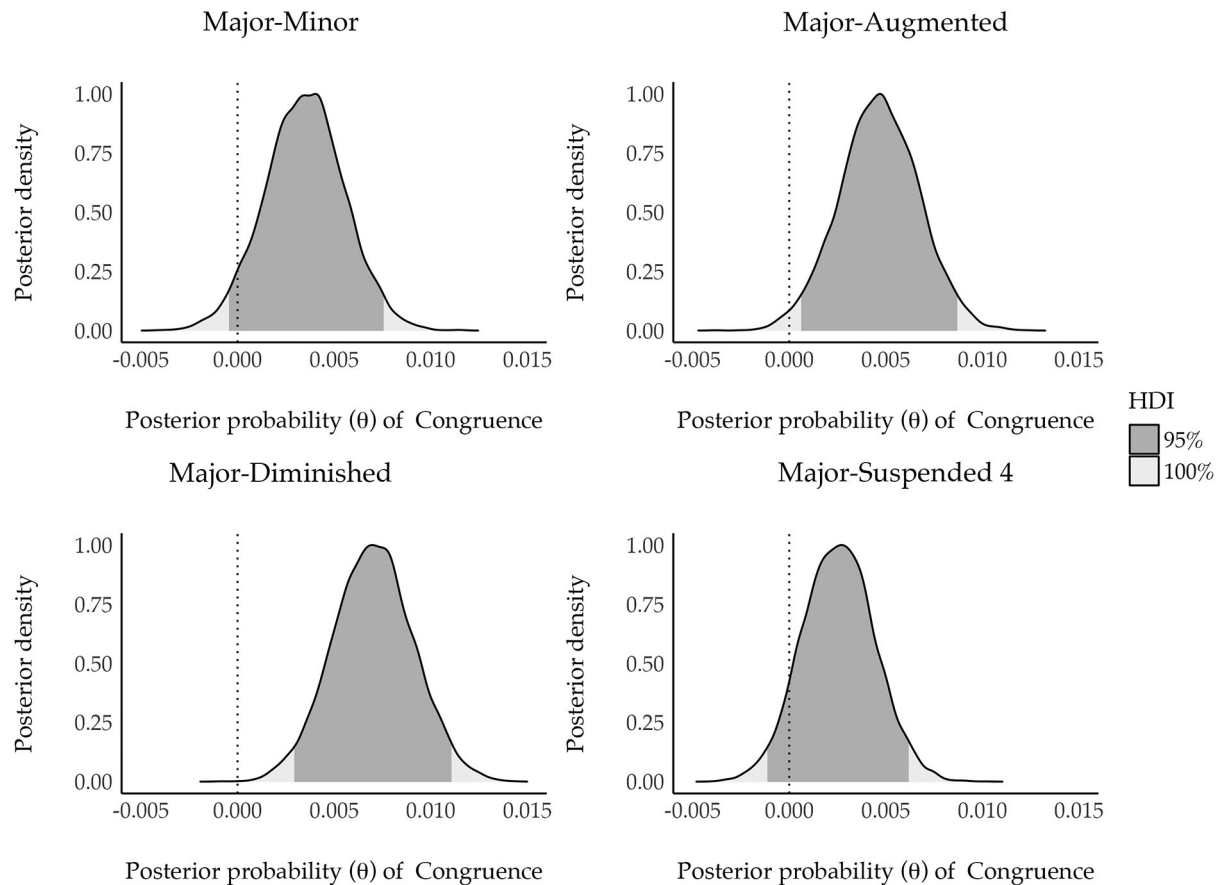


FIGURE 2. Posterior probability distributions of the Congruence factor for the four sub-experiments. The grey area marks the 95% credibility interval and the vertical line highlights the location of θ zero.

greater than the congruence for the augmented chord. However, this is not a sufficiently large difference if we apply the credible interval to this delta posterior distribution, which shows that 95% delta posterior distribution contains zero (-0.003–0.008 with the median of 0.0024). Our interpretation about the difference is that while the diminished chord has a higher probability of receiving a larger affective priming effect than the augmented chord, this difference is not substantial.

Discussion

The present study examined whether acquired connotations of common tonal chords could give rise to affective priming effects when chords with putatively negative associations (minor, augmented, and diminished) are paired with a major chord (positive associations) as primes with negative and positive word targets. Three

out of four hypotheses were corroborated by the analysis and the notion that affective priming can be achieved without the involvement of roughness was strongly supported. The only chord pairing that failed to receive support was Major – Minor. We did not expect the suspended fourth chord to be able to generate negative associations and the results indeed supported this proposition. This sub-experiment acted as the general control condition of the experiment and operated as predicted. In sum, the affective priming results were notably in line with the self-reported valence ratings in previous research (Bowling et al., 2018; Lahdelma & Eerola, 2016a; Roberts, 1986) in that only the relatively unpleasant sounding chords (namely the augmented and diminished chords) created clear negative congruence effects. Next, we will discuss the results in terms of each individual chord pair and their respective hypotheses that were put forward in the pre-registered report.

Major-Minor: Somewhat surprisingly the minor chord worked less effectively as a prime than hypothesized. This finding, however, is reconcilable with previous research if we take into account the role of musical sophistication. Using nonmusicians as participants, Costa (2013) found that the minor chord was not an effective prime when using sad words as targets, and Steinbeis and Koelsch (2011) similarly found that this mode of perception was more robust in the case of musicians when compared to nonmusicians. Interestingly, Costa (2013) also found that valenced images (instead of words) did not create any congruence effects with major/minor chords as primes. In terms of nonmusicians, the only report of a robust effect so far is from Bakker and Martin (2015) who reported that when valenced faces and major/minor chords were presented simultaneously, processing was facilitated, as indexed by decreased N2 ERP amplitudes (understood as facilitation of early processing). The evident differences in modes of presentation (stimulus onset asynchrony vs. simultaneous presentation, faces instead of words/pictures), the data analysis approach (using averaged data instead of modeling raw responses), as well as the role of musical expertise across these presented studies all warrant further investigation based on the current results. The lack of strong evidence for the major/minor distinction among musically naïve listeners, however, is in line with developmental studies (Dalla Bella et al., 2001; Gregory et al., 1996) and cross-cultural data (Athanasopoulos et al., 2021; Lahdelma et al., 2021; Smit et al., 2022) where this mode of perception has been linked to learning and enculturation instead of an alleged inherent affective response arising from similarity to positive/negative speech (cf. Bowling et al., 2012; Curtis & Bharucha, 2010).

Major-Augmented: this chord pair provided strong affective priming results as expected, in line with the low pleasantness ratings obtained through self-report data in previous experiments (Bowling et al., 2018; Lahdelma & Eerola, 2016a; Roberts, 1986). Also, music historical observations suggest that this chord has been used to denote negative valence since the 19th century onwards (Huckvale, 1990; Todd, 1988). As the chord's negative valence is hard to pinpoint acoustically in terms of roughness/harmonicity (see Table 1), and since the perception of negative valence in this chord seems to be absent in non-Western cultures (Lahdelma et al., 2021; McDermott et al., 2016), the remaining explanation for the augmented chord's negative valence seems indeed to be enculturation.

Major-Diminished: this chord pair provided strong affective priming results as expected, in line with the

low pleasantness ratings obtained through self-report data in previous experiments (Bowling et al., 2018; Lahdelma & Eerola, 2016a) much like in the augmented chords's case. The possible contribution of inharmonicity cannot be ruled out at this stage as the diminished chord is somewhat inharmonic (in the 50%–75% quantile in a sample of chord harmonicities, see Table 1), but a competing cultural interpretation is also a plausible explanation. It is possible that the diminished chord represents simply a more firmly consolidated convention than the augmented chord: after all, the diminished chord is part of major and (harmonic) minor key harmonisations on the VII degree and (natural and harmonic) minor key harmonisations on the II degree, whereas the augmented chord is only present in harmonic minor keys (on the III degree). As major keys are significantly more common than minor keys in Western tonal music and represent a de facto *norm* (Parncutt, 2014), a dissonant chord used in both major and minor keys is more familiar than a dissonant chord used only in (harmonic) minor keys (see also Johnson-Laird et al., 2012). The diminished chord has been a prominent chord since the Renaissance and became even more prevalent in the Baroque period due to being perceived as incomplete dominant sevenths (Parncutt & Hair, 2011). The diminished chord's negative affective convention goes back further in music history than the augmented chord's: the latter became a clearly independent sonority only in the 19th century (Todd, 1988), whereas the diminished chord's affective use goes back to at least the Baroque period: according to Hatten (2004) there is a straightforward expressive association between the diminished (including also the seventh note) chord and a sense of tension, “or more specifically, human angst, that has a long history of rhetorical usage going back to early Baroque recitative” (p. 49). By the 18th century the diminished chord was already established as a “cliché” of horror and terror (Garlington, 1963); this connotation was only reinforced later in film music where it was equally applied as a “horror chord” starting with early silent movies (Tagg, 2014). As opposed to the augmented chord, there is no cross-cultural data on how the diminished chord is perceived in non-Western cultures; McDermott et al. (2016) did collect data on this chord in their study but did not report the results on the basis that in their stimuli the major and augmented chords produced the largest contrast in pleasantness ratings in US listeners; on the basis of some similar self-report studies the contrasts between the major and diminished chords are indeed less pronounced than the major and augmented chords and this difference seems to be related to musical sophistication

(see data in e.g., Arthurs et al., 2018; Roberts, 1986), although other studies have found a more closely matched difference between the major chord and the augmented and diminished chords in terms of valence when collapsing data from both musicians and nonmusicians (Bowling et al., 2018; Lahdelma & Eerola, 2016a).

Major-Suspended Fourth: as expected, this chord pair did not create affective priming results, in line with the high pleasantness ratings obtained through self-report data in previous experiments (Arthurs et al., 2018; Bowling et al., 2018). The notable element in the suspended chord's inability to prime negative valence is that the chord contains a major second interval that has been shown to create automatic negative congruency alongside the minor second interval in a previous affective priming study (Armitage et al., 2021). This finding has interesting implications for research on automatic reactions to harmony: it seems that additional acoustic information in the suspended fourth chord is enough to "overrule" the major second interval's roughness so that no automatic aversion is observed in response to this chord despite the presence of the rough major second interval, and this observation is notably in line with the surprisingly low global roughness value of the chord. Several theoretical possibilities emerge to explain this finding. 1) Wright and Bregman (1987) have suggested that perceived roughness in a single chord may be mitigated by placing it in a (in this case hypothetical) contrapuntal context. As the suspended fourth contains tension stemming from an implied voice leading situation in the 4th degree's need of resolution to either the minor 3rd or more typically the major 3rd (Parncutt et al., 2019), it is possible that this tonal tendency indeed mitigates the sensory roughness of the major second interval within the chord through this implied motion. As nonmusicians have been demonstrated to be sensitive to tension in terms of both single isolated chords (Lahdelma & Eerola, 2020) and chord progressions (Bigand & Parncutt, 1999) this possibility warrants further investigation. 2) It is possible that the major second interval's roughness is simply irrelevant in a chord that contains positive valence through associative learning: as the chord most typically resolves to a major chord in tonal music (Parncutt et al., 2019), it is possible that its positive valence is learned through this contextual association. 3) Finally, there exists a possibility that there is some interaction with the combined overtones of the individual fundamentals present in the suspended fourth chord that effectively mitigates the sensory roughness of the major second

interval through perceptual fusion (for similar observations about the major triad "overruling" the harshness of the minor second interval within the major seventh chord, see Mashinter, 2006, and Lahdelma & Eerola, 2015). What can already be concluded from the suspended fourth chord's case is that the arising quality of a chord can evidently be quite different from the sum of its parts, as per the classic Gestalt notion of holistic perception (Langfeldt, 2022); this finding implies that a holistic perception of chords can surpass their raw acoustic qualities (e.g., the roughness of specific intervals that is not perceived due to other acoustic information present in the chord).

In terms of the current study's limitations, the participant pool and the variety of stimuli could be further diversified in future research. Here we on purpose focused on nonmusicians to investigate how the general population familiar with Western music perceives common tonal chords when measured with an affective priming method, and the stimuli were hence restricted to some of the most fundamental chords used in Western music. In the future, it would be also crucial to corroborate the findings obtained via the affective priming method with other indirect and implicit measures such as EEG/ERP (Eder et al., 2012) or pupillometry responses (Laeng et al., 2012).

Based on prior literature, we hypothesized that the minor, the diminished, and the augmented chords would elicit affective priming when paired with the generally positive prime (major chord). Indeed, the augmented and diminished chords clearly produced the assumed effect although the minor chord fell short of a robust effect. We demonstrated that affective priming is possible without extreme contrasts in roughness/harmonicity, and that these effects cannot be attributed to chord frequency per se. In the current experiments, the results of the automatic responses were notably in line with perceived valence, which in turn seem to reflect the influence of enculturation to Western common practice music. The unexpected failure to observe a difference in affective priming between the major and minor chords may be due to our largely nonmusician sample since previous research (Costa, 2013; Steinbeis & Koelsch, 2011) has correspondingly also found weaker affective priming differences between major and minor chords in nonmusicians compared to musicians.

On the basis of our results, we recommend further research to investigate the role of expertise, as well as the holistic perception of chords that may mitigate the aversion to roughness in certain pitch combinations when measured with an automatic reaction method.

Author Note

Correspondence concerning this article should be addressed to Imre Lahdelma, Durham University,

Department of Music, Palace Green, DH1 3RL, Durham, United Kingdom. E-mail: imre.d.lahdelma@durham.ac.uk

References

- ARMITAGE, J., LAHDELMA, I., & EEROLA, T. (2021). Automatic responses to musical intervals: Contrasts in acoustic roughness predict affective priming in western listeners. *Journal of the Acoustical Society of America*, *150*(1), 551–560.
- ARTHURS, Y., BEESTON, A. V., & TIMMERS, R. (2018). Perception of isolated chords: Examining frequency of occurrence, instrumental timbre, acoustic descriptors and musical training. *Psychology of Music*, *46*(5), 662–681.
- ATHANASOPOULOS, G., EEROLA, T., LAHDELMA, I., KALIAKATSOS-PAPAKOSTAS, M., & DAMJANOVIC, L. (2021). Harmonic organisation conveys both universal and culture-specific cues for emotional expression in music. *PLOS ONE*, *16*(1), e0244964.
- BAKKER, D. R., & MARTIN, F. H. (2015). Musical chords and emotion: Major and minor triads are processed for emotion. *Cognitive, Affective, and Behavioral Neuroscience*, *15*(1), 15–31.
- BARR, D. J., LEVY, R., SCHEEPERS, C., & TILY, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.
- BATES, D., MÄCHLER, M., BOLKER, B., & WALKER, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- BAYESIAN REPORTING GUIDELINES. (n.d.). Retrieved from easystats.github.io/bayestestR
- BIGAND, E., & PARNCUTT, R. (1999). Perceiving musical tension in long chord sequences. *Psychological Research*, *62*(4), 237–254.
- BIGAND, E., PARNCUTT, R., & LERDAHL, F. (1996). Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception and Psychophysics*, *58*(1), 125–141.
- BOWLING, D., PURVES, D., & GILL, K. Z. (2018). Vocal similarity predicts the relative attraction of musical chords. *Proceedings of the National Academy of Sciences*, *115*(1), 216–221.
- BOWLING, D., SUNDARARAJAN, J., HAN, S., & PURVES, D. (2012). Expression of emotion in Eastern and Western music mirrors vocalization. *PLoS One*, *7*(3), e31942.
- BROYLES, J. C. (1999). *How dissonant is the augmented triad?* [PhD thesis, University of British Columbia].
- BROZE, Y., & SHANAHAN, D. (2013). Diachronic changes in jazz harmony: A cognitive perspective. *Music Perception*, *31*(1), 32–45.
- BRYLSBAERT, M., & STEVENS, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, *1*(1). <https://doi.org/10.5334/joc.10>
- BURGOYNE, J. A. (2012). *Stochastic processes and database-driven musicology* [PhD thesis, McGill University].
- BÜRKNER, P.-C. (2021). Bayesian item response modeling in R with brms and Stan. *Journal of Statistical Software*, *100*(5), 1–54.
- COHEN, A. (2001). Music as a source of emotion in film. In P. Juslin & J. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 879–908). Oxford University Press.
- COSTA, M. (2013). Effects of mode, consonance, and register in visual and word-evaluation affective priming experiments. *Psychology of Music*, *41*(6), 713–728.
- CROWDER, R. G., REZNICK, J. S., & ROSENKRANTZ, S. L. (1991). Perception of the major/minor distinction: V. Preferences among infants. *Bulletin of the Psychonomic Society*, *29*(3), 187–188.
- CURTIS, M. E., & BHARUCHA, J. J. (2010). The minor third communicates sadness in speech, mirroring its use in music. *Emotion*, *10*(3), 335–348.
- DALLA BELLA, S., PERETZ, I., ROUSSEAU, L., & GOSSELIN, N. (2001). A developmental study of the affective value of tempo and mode in music. *Cognition*, *80*(3), B1–10.
- DE CLERCQ, T., & TEMPERLEY, D. (2011). A corpus analysis of rock harmony. *Popular Music*, *30*(1), 47–70.
- DE HOUWER, J., TEIGE-MOCIGEMBA, S., SPRUYT, A., & MOORS, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin*, *135*(3), 347–368.
- EDER, A. B., LEUTHOLD, H., ROTHERMUND, K., & SCHWEINBERGER, S. R. (2012). Automatic response activation in sequential affective priming: An ERP study. *Social Cognitive and Affective Neuroscience*, *7*(4), 436–445.
- EEROLA, T., & LAHDELMA, I. (2021). The anatomy of consonance/dissonance: Evaluating acoustic and cultural predictors across multiple datasets with chords. *Music and Science*, *4*, 1–19.
- EEROLA, T., & LAHDELMA, I. (2022). Register impacts perceptual consonance through roughness and sharpness. *Psychonomic Bulletin and Review*, *29*(3), 800–808.
- FAZIO, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition and Emotion*, *15*(2), 115–141.

- FRIEDMAN, R. S., KOWALEWSKI, D. A., VUVAN, D. T., & NEILL, W. T. (2021). Consonance preferences within an unconventional tuning system. *Music Perception*, 38(3), 313–330.
- GAO, C., WEDELL, D. H., & SHINKAREVA, S. V. (2020). Crossmodal negativity bias in semantic processing. *Emotion*, 22(6), 1270–1280.
- GARLINGTON, A. S. (1963). “Le Merveilleux” and operatic reform in 18th-century French opera. *The Musical Quarterly*, 49(4), 484–497.
- GERARDI, G. M., & GERKEN, L. (1995). The development of affective responses to modality and melodic contour. *Music Perception*, 12(3), 279–290.
- GREGORY, A. H., WORRALL, L., & SARGE, A. (1996). The development of emotional responses to music in young children. *Motivation and Emotion*, 20, 341–348.
- HARDIE, S. M., & WRIGHT, L. (2014). Differences between left- and right-handers in approach/avoidance motivation: Influence of consistency of handedness measures. *Frontiers in Psychology*, 5(134), 1–10.
- HARRISON, P., & PEARCE, M. (2020). Simultaneous consonance in music perception and composition. *Psychological Review*, 127(2), 216–244.
- HATTEN, R. S. (2004). *Musical meaning in Beethoven: Markedness, correlation, and interpretation*. Indiana University Press.
- HERRING, D. R., WHITE, K. R., JABEEN, L. N., HINOJOS, M., TERRAZAS, G., REYES, S. M., ET AL. (2013). On the automatic activation of attitudes: A quarter century of evaluative priming research. *Psychological Bulletin*, 139(5), 1062–1089.
- HUCKVALE, D. (1990). Twins of evil: An investigation into the aesthetics of film music. *Popular Music*, 9(1), 1–35.
- HURON, D. (2001). Tone and voice: A derivation of the rules of voice-leading from perceptual principles. *Music Perception*, 19(1), 1–64.
- HUTCHINSON, W., & KNOPOFF, L. (1978). The acoustic component of Western consonance. *Interface*, 7(1), 1–29.
- JOHNSON-LAIRD, P. N., KANG, O. E., & LEONG, Y. C. (2012). On musical dissonance. *Music Perception*, 30(1), 19–35.
- KIVY, P. (2002). *Introduction to a philosophy of music*. Oxford University Press.
- KRUSCHKE, J. K. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. Academic Press.
- KRUSCHKE, J. K. (2018). Rejecting or accepting parameter values in bayesian estimation. *Advances in Methods and Practices in Psychological Science*, 1(2), 270–280.
- LAENG, B., SIROIS, S., & GREDEBÄCK, G. (2012). Pupillometry: A window to the preconscious? *Perspectives on Psychological Science*, 7(1), 18–27.
- LAHDELMA, I., ARMITAGE, J., & EEROLA, T. (2022). Affective priming with musical chords is influenced by pitch numerosity. *Musicae Scientiae*, 26(1), 208–217.
- LAHDELMA, I., ATHANASOPOULOS, G., & EEROLA, T. (2021). Sweetness is in the ear of the beholder: Chord preference across United Kingdom and Pakistani listeners. *Annals of the New York Academy of Sciences*, 1502(1), 72–84.
- LAHDELMA, I., & EEROLA, T. (2015). Theoretical proposals how vertical harmony may convey nostalgia and longing in music. *Empirical Musicology Review*, 10(3), 245–263. <https://doi.org/http://emusicology.org/article/view/4534/4151>
- LAHDELMA, I., & EEROLA, T. (2016a). Single chords convey distinct emotional qualities to both naive and expert listeners. *Psychology of Music*, 44(1), 37–54.
- LAHDELMA, I., & EEROLA, T. (2016b). Mild dissonance preferred over consonance in single chord perception. *I-Perception*, 7(3), 1–21.
- LAHDELMA, I., & EEROLA, T. (2020). Cultural familiarity and musical expertise impact the pleasantness of consonance/dissonance but not its perceived tension. *Scientific Reports*, 10, 8693.
- LAHDELMA, I., EEROLA, T., & ARMITAGE, J. (2022). Is harmonic a misnomer for cultural familiarity in consonance preferences? *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.802385>
- LAKENS, D., & CALDWELL, A. (2021). Simulation-based power analysis for factorial analysis of variance designs. *Advances in Methods and Practices in Psychological Science*, 4(1), 251524592095150. <https://doi.org/10.1177/2515245920951503>
- LANGFELDT, V. (2022). *Holistic identification of musical harmony. A theoretical and empirical study* [PhD thesis, Norwegian Academy of Music].
- LINNAVALLI, T., OJALA, J., HAVERI, L., PUTKINEN, V., KOSTILAINEN, K., SEPPÄNEN, S., & TERVANIEMI, M. (2020). Musical expertise facilitates dissonance detection on behavioral, not on early sensory level. *Music Perception*, 38(1), 78–98.
- MAKOWSKI, D., BEN-SHACHAR, M. S., CHEN, S. A., & LÜDECKE, D. (2019). Indices of effect existence and significance in the bayesian framework. *Frontiers in Psychology*, 10, 2767.
- MASHINTER, K. (2006). Calculating sensory dissonance: Some discrepancies arising from the models of Kameoka & Kuriyagawa, and Hutchinson & Knopoff. *Empirical Musicology Review*, 1, 65–84.
- MCDERMOTT, J. H., SCHULTZ, A. F., UNDURRAGA, E. A., & GODOY, R. A. (2016). Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature*, 535(7613), 547–550.
- MEYER, L. B. (1956). *Emotion and meaning in music*. Chicago University Press.
- OLLEN, J. E. (2006). *A criterion-related validity test of selected indicators of musical sophistication using expert ratings* [PhD thesis, The Ohio State University].
- PARNCUTT, R. (1989). *Harmony: A psychoacoustical approach*. Springer-Verlag.

