Automatic Responses to Acoustically Rough Musical Intervals

James Armitage, Imre Lahdelma, and Tuomas Eerola

Department of Music, Durham University, Durham, DH1 3RL, UK

(Dated: 25 August 2020)

The aim of the present study is to determine which acoustic components of harmonic consonance and dissonance influence automatic responses in a simple cognitive task. In a series of experiments, ten musical interval pairs were used to measure the influence of acoustic roughness and harmonicity on response times in an affective priming task conducted online. There was a significant correlation between the difference of roughness for each pair of intervals and a response time index. Harmonicity did not influence response times on the cognitive task. More detailed analysis suggests that the presence of priming in intervals is binary: in the negative primes that create congruency effects the intervals' fundamentals and overtones coincide within the same equivalent rectangular bandwidth (i.e. the minor and major seconds). Intervals that fall outside this equivalent rectangular bandwidth do not elicit priming effects, regardless of their dissonance or cultural conventions of negative affect. The results are discussed in the context of recent developments in consonance/dissonance research and vocal similarity.

I. INTRODUCTION

The contrast between consonance and dissonance is a vital feature of Western music. Consonance is typically perceived as *agreeable* and *stable* while dissonance, in turn, as *disagreeable* and *in need of resolution* (Tramo *et al.*, 2001). Consonance/dissonance has both a vertical and a horizontal aspect: single isolated *intervals* (two concurrent pitches) and *chords* (three or more concurrent pitches) represent *vertical consonance/dissonance*, while the sequential relationships between these in melodies and chord progressions represent *horizontal consonance/dissonance* (Parncutt and Hair, 2011). The current research refers exclusively to the vertical aspect.

Empirical research concerning consonance and dissonance frequently relies on self-report methods, which come with a well-documented set of limitations (see e.g. Fazio and Olson, 2003, for review). Priming on the other hand captures participants' automatic responses to the stimuli, avoiding demand characteristics and tapping a construct that is underpinned by both physiology and culture (see e.g. Herring *et al.*, 2013, for review). Previous studies using an affective priming paradigm have shown that valenced chords (e.g. consonant-positive, dissonant-negative) facilitate the evaluation of similarly valenced target words (see e.g. Steinbeis and Koelsch, 2011). Recent research (Lahdelma *et al.*, 2020) has found that this congruency effect is not present when intervals, as opposed to chords, are used as primes. This suggests that the findings are related to acoustic components such as *roughness* and *harmonicity*. Roughness is often seen as prevalent in the perception of dissonance but not in the perception of consonance (see e.g. Hutchinson and Knopoff, 1979), as dissonant intervals contain less overall roughness than dissonant chords. The study by Lahdelma *et al.* (2020), however, tested only four distinct intervals and excluded some of the most culturally loaded ones in terms of positive and negative affect (e.g. the major/minor thirds and the tritone), which poses a possible limitation in disentangling the acoustic variable from the cultural. Another major acoustic factor related to dissonance is *harmonicity*, which has been demonstrated to contribute to perception of consonance in a variety of settings (McDermott *et al.*, 2010). Finally, support for a culturally-based argument is perhaps provided by a study conducted by Steinbeis and Koelsch (Steinbeis and Koelsch, 2011) which found that congruency effects could be primed by major and minor triads which lack clear differences in roughness.

The present study aims to disentangle the role of specific acoustic properties of the intervals, namely *roughness* and *harmonicity* and consider contrasts that stem from *affective conventions*, such as the contrasts between major and minor thirds and sixths, rather than acoustics.

A. Consonance and Dissonance

1. Acoustic Roughness

Historically, the acoustic and perceptual characteristics of consonance and dissonance were placed on a sound theoretical and empirical footing by von Helmholtz (1885) who proposed acoustic roughness as an explanation for why some musical intervals are considered dissonant and disagreeable. There is consensus that the sensation of roughness is caused by interference patterns between wave components of similar frequency that gives rise to beating (see e.g. Hutchinson and Knopoff, 1978), which in turn creates the sound qualitys of that listeners typically perceive as unpleasant. Based on his findings of the relationship between roughness and dissonance, von Helmholtz derived that consonance in turn is the absence of roughness, and concluded that roughness is the cause of both consonance and dissonance in music. Coming to a similar conclusion later, Terhardt (1984) proposed that the evaluation of consonance in isolated intervals and chords is mostly governed by *sensory consonance*, i.e. a lack of unpleasant features of a sound such as sharpness (the presence of spectral energy at high frequencies) and roughness.

The perception of roughness has a biological substrate, as beating occurs at the level of the basilar membrane in the inner ear when the frequency components are too close together to separate (see e.g. Peretz, 2010). This range is known as the *critical bandwidth* (Fletcher and Munson, 1933). According to Smith and Abel (1999, p. 21) ”a critical band is 100 Hz wide for center frequencies below 500 Hz, and 20% of the center frequency above 500 Hz”. A more recent formulation of this concept, the equivalent rectangular bandwidth (ERB, see Patterson, 1976) approximates the basilar membrane as being made up of rectangular band-pass filters. Fundamental frequencies processed within the same band-pass filter are perceived as acoustically rough. Compared to the critical bandwidth, the ERB typically encompasses a narrower range of frequencies (Smith and Abel).

The sensitivity to roughness seems to be present cross-culturally (see McDermott *et al.*, 2016), but its appraisal differs significantly across musical styles and cultures: while a typical Western listener hears roughness as disagreeable, it is deliberately harnessed in the vocal practice of *beat diaphony* (known as *Schwebungsdiaphonie* in German literature) in for example the Baltic and Balkan regions of Europe (see e.g. Ambrazevičius, 2017) and in Papua New Guinea (Messner, 1981). Several models of roughness exist such as those by Hutchinson and Knopoff (1978), Vassilakis (2001), and Wang *et al.* (2013). These models, which based on emulation human auditory system in sensation of roughness, largely agree on the amount of roughness in different intervals. Figure 1 shows how roughness varies for intervals over the octave from C_4 to C_5 as the distance between the two fundamentals increases by one cent – equal to 1/1200 of an octave.

While roughness was long accepted as a sufficient explanation for consonance and dissonance (see Helmholtz, 1885; Terhardt, 1984), later counterarguments were made for why it is not all-encompassing in explaining its underlying cause. First, perceptions of consonance and dissonance seem to remain when the tones of a chord are presented independently to the ears (i.e. dichotically), precluding physical interaction at the input stage and thus greatly reducing the perception of roughness (Bowling and Purves, 2015; McDermott *et al.*, 2010). However, it has been pointed out that beats could also occur centrally, within a binaural critical band rather than being based on cochlear interactions (Carcagno *et al.*, 2019). Second, the perceived consonance of a chord does not seem to increase when roughness is artificially removed (Nordmark and Fahlén, 1988). Third, it has been shown that participants with congenital amusia (i.e. a neurogenetic disorder characterised by an inability to recognise or reproduce musical tones) exhibit abnormal consonance perception but normal roughness perception (Cousineau *et al.*, 2012).

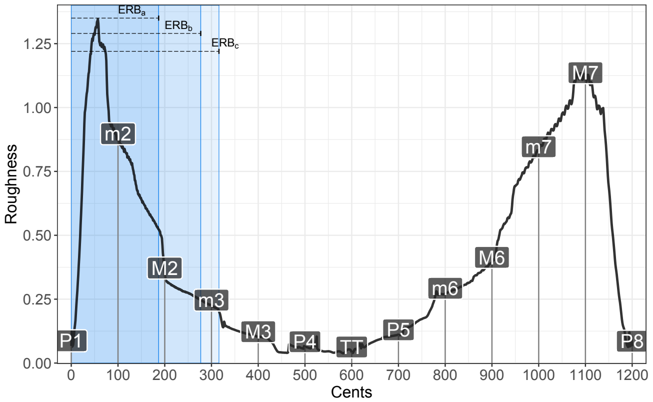

FIG. 1. Roughness using the model by Wang *et al.* (2013) for intervals from unison to octave divided into 1200 cents. Frequency-dependent critical bandwidth boundaries are shown for the first three intervals ($\mathrm{ERB_a} = \mathrm{m2}$, $\mathrm{ERB_b} = \mathrm{M2}$, $\mathrm{ERB_c} = \mathrm{m3}$) reflecting the different mean frequencies of the intervals in our design (see Stimuli) using the ERB bandwidths (Moore and Glasberg, 1983).

2. Harmonicity

Another possible explanation that has been put forward to explain consonance and dissonance in addition to roughness is the acoustic property of *harmonicity* which denotes how closely the spectral frequencies of a sound correspond to a harmonic series. A single musical pitch is the combination of the fundamental frequency plus its overtones, which typically follow the harmonic series. If an interval is made of two tones which share a large proportion of their overtones (e.g. the perfect fifth) then the interval is high in harmonicity. Conversely, intervals whose overtones do not significantly overlap (e.g. the major seventh) are considered low in harmonicity. In other words, harmonicity posits that in consonant pitch combinations the component frequencies produce an aggregate spectrum that is typically harmonic i.e.

it resembles the spectrum of a single sound. In contrast, dissonant intervals and chords produce an inharmonic spectrum (Cousineau *et al.*, 2012).

Harmonicity offers a conception of consonance that is defined constructively rather than by the absence of roughness. Like roughness, harmonicity may have a biological underpinning, albeit at the neuronal rather than auditory level (Tramo *et al.*, 2001), even if this is debated (Carcagno *et al.*, 2019). Preference for high levels of harmonicity has been found to correlate with preference for consonance (McDermott *et al.*, 2010), and it has been suggested that this link might be related to the advantages of recognising human vocalisation (Bowling and Purves, 2015). Another possibility is that, unless humans are born with a preference for harmonicity, it is acquired simply through exposure to natural sounds or to music (McDermott *et al.*, 2010).

3. Affective Conventions

Apart from consonance and dissonance *per se*, certain musical intervals also bear strong affective conventions in Western culture. The minor third is typically associated with sadness while the major third with positive affect (Curtis and Bharucha, 2010). This is mirrored also in speech cues where the spectra of major intervals are more similar to spectra found in excited speech, and the spectra of minor intervals, in turn, are more similar to the spectra of subdued speech (Bowling *et al.*, 2010). The tritone carries conventionally a negative connotation in Western music (see e.g. Costa *et al.*, 2000) and has historically been described as 'diabolic' (Partch, 1974). In an empirical setting, minor intervals have been found to be perceived as more 'gloomy' and 'sinister' compared to their major counterparts (Costa *et al.*, 2000), and the major third have been demonstrated to be associated with adjectives like 'calm', 'pleasing', and 'happy' (Oelmann and Laeng, 2009). Also, the major sixth has been found to convey 'joy' to listeners while the major seventh, in turn, 'sadness' (Krantz *et al.*, 2004).

B. The Present Study

The extent to which acoustical aspects (roughness, harmonicity) contribute to the perception of consonance/dissonance in intervals has remained contentious. The present study seeks to explore the extent to which harmonicity and roughness influence automatic responses on a simple cognitive task. As affective priming is an indirect measure, it is expected to yield valuable information on the importance of these individual factors' contributions to the perception of consonance and dissonance. The present study considers whether intervals can influence behaviour on a word evaluation task. An important question that has remained unclear in previous research is what property of the prime is responsible for activating the nodes in the semantic-affective network that lead to affective priming – perceived valence or consonance/dissonance (teasing apart the components roughness and harmonicity). The present study uses ten interval pairs which vary in terms of valence (e.g. major and minor thirds) and consonance/dissonance (e.g. minor second and perfect fifth).

In a similar previous experiment (Lahdelma *et al.*, 2020), it was found that when valenced words were preceded by tetrachords (four concurrent pitches), there were significant congruency effects – i.e. positive words were classified more quickly when preceded by a consonant rather than a dissonant chord, and vice versa for negative words. Notably, these congruency effects were absent for intervals. To address the uncertainty around context discussed in Lahdelma *et al.*, the present experiments set out to examine whether diatonic intervals can drive priming effects without the confound of exposure to more complex chords.

In ten separate sub-experiments conducted online, one for each pair of intervals (see Table II), participants were asked to classify a sequence of emotional words as either positive or negative. Each word was preceded by the brief sounding of a musical interval. To test whether intervals carrying cultural conventions of positive/negative affect can elicit congruency effects with similarly valenced words, the major and minor thirds and the major and minor sixths were tested, as well as the tritone. Additionally, we evaluated the extent to which manipulating the amount of roughness can influence results of the behavioural task. To this end we tested two interval pairs involving artificially detuned minor seconds (played with both the piano timbre and with the Shepard tone, see Materials & Stimuli for details) in order to maximise roughness. Preferential response to acoustic roughness is thought to confer an evolutionary advantage; for instance alarm signals, whether in nature or man-made, are frequently high in roughness (Arnal *et al.*, 2015). Indeed, specifically in the case of human vocalisation, roughness has been linked to perceived anger (Bänziger *et al.*, 2015), screams (Schwartz *et al.*, 2019), and infant cries (Koutseff *et al.*, 2018). Consequently, it was predicted that intervals which are high in roughness will prime responses to negative stimuli more effectively than those which are typically perceived as negative in valence, such as the minor third (Curtis and Bharucha, 2010), but that are low in roughness. More specifically, we tested the hypothesis that priming effects are influenced by differences in the acoustic variables roughness and harmonicity associated with the intervals in each pair of primes: details of the differences are provided in Table II in the Materials & Stimuli section.

II. METHODS

A. Participants

Participants were recruited via Prolific Academic (`https://prolific.ac`), a crowdsourcing platform targeted specifically for academic research. Following deletions (17 participants; see Data Analysis), 387 participants (197 female, 181 male, 3 other, mean age = 36.2, SD = 12.4) completed the study. All participants reported corrected to normal or normal vision and were right-handed native speakers of English. 295 participants identified as non-musicians. Ethical approval was granted by the Music Department Ethics Committee, University of anonymous for review. Informed consent was provided via an online checkbox.

B. Materials & Stimuli

Ten prime stimuli were generated in total, eight diatonic intervals (equal-tempered) plus two intervals manipulated in tuning to maximise roughness. The diatonic intervals were created in accordance with the procedure employed by Bowling et al. (Bowling *et al.*, 2018) where the fundamental frequencies (f_0) of the pitches in each interval were adjusted so that the mean f_0 of all pitches was C_4 (261.63 Hz). Descriptors of roughness and harmonicity are given in Table I. Roughness calculations were carried out using the model developed by Wang *et al.* (2013). The analyses were duplicated using the Vassilakis (2001) and Hutchinson and Knopoff (1978) models, and with a composite model (mean roughness value of all three models). The choice of roughness model did not alter the pattern of significance of results. Harmonicity was calculated using the model by Harrison and Pearce (2018) which simulates the way listeners search the auditory spectrum for occurrences of harmonic spectra; harmonicity values under alternative harmonicity models and a composite model are detailed in the supplementary material.

As an additional diagnostic measure of dissonance, the first thirteen partials (fundamental plus twelve overtones) were extracted from each single tone from the intervals presented below. For each interval, we calculated the ERB about the mean of each pair of partials, generating 169 (i.e. 13 × 13) ERBs and checked whether the frequencies fell within this band. To determine which partials fell within the ERBs, we used the formula derived by Moore and Glasberg (1983) to calculate the width of an ERB centered on the mean frequency of each partial pair:

$$ERB = 6.23f^2 + 93.3f + 28.52$$

Table I details how many partial pairs for each interval fell within ERBs. The ERB was defined as the frequency band between the boundary suggested by Moore and Glasberg (Moore and Glasberg, 1983) and 10 Hz (beating effects due to frequency differences of less than 10 Hz would be perceived as amplitude modulation or 'beats' rather than roughness (see e.g. Roederer, 2008, p. 38); a further caveat is that very close alignments of overtones would contribute to a sense of harmonicity, for instance in the case of P5).

TABLE I. Intervals, Notation, and Key Descriptors.

Interval (Abbr.)	Roughness	Harmonicity	Partials/ERB
Minor second (m2)	10.00	1.00	23
Major second (M2)	4.79	5.41	25
Minor third (m3)	3.66	2.80	18
Major third (M3)	4.66	4.72	15
Perfect fourth (P4)	3.64	10.00	14
Tritone (TT)	2.70	1.03	15
Perfect fifth (P5)	1.00	10.00	13
Minor sixth (m6)	4.75	4.72	14
Major sixth (M6)	5.98	2.80	15
Minor seventh (m7)	5.70	5.41	14
Major seventh (M7)	2.58	1.00	12

For the primes, eight intervals were combined into ten different pairings: m3/M3, m6/M6, TT/P5, m2/P5, and M7/P5, m2/TT, M2/P5, m2/M3, d2/P5 (detuned minor second) and s2/S5 (Shepard tones). The first eight interval pairings were generated with *Ableton Live 9* (a music sequencer software), using the *Synthogy Ivory Grand Pianos II* plug-in. For the piano interval pairs, the applied sound font was *Steinway D Concert Grand.* No reverb was used, and the intervals had a fixed velocity (65) in order to have a neutral and even sound. In the d2/P5 interval pairing the minor second interval was created by taking a unison and detuning one pitch down by -90 cents; this procedure created a notably high amount of roughness when measured with the models by Vassilakis (2001) and Wang *et al.* (2013). The pairing s2/S5 was created using Shepard tones that have octave-spaced partials from 16Hz to 20kHz with cosine curve shaped spectral envelope. The detuning of the minor second interval using the Shepard tone was created by shifting the odd partials above the fundamental upward and partials below downward by a detuning constant d. The constant was determined to yield a maximal roughness at $d = 0.024$ using the roughness models by Vassilakis and Wang *et al.*. Table II shows the pairs of primes and the differences in their acoustic parameters. The classification as High or Low ΔRoughness and High or Low ΔHarmonicity remained unchanged when we calculated composite Roughness and Harmonicity measures (see Supplementary Material).

The loudness of the stimuli was equalised by setting them to the same peak sone level. The sound files were converted to stereo (same signal in both channels) as 44.1 kHz, 32 bits per sample waveform audio files. The length of each interval was exactly 800 ms including a 10 ms fadeout.

The target words were chosen from the database compiled by Warriner et al. (Warriner *et al.*, 2013): *Flaccid, Hijack, Climax, Gentle, Lively, Rest, Excite, Payday, Rabid, Coma, Saggy, Dismal, Relax, Comfy, Arrest, Morgue*. The target words were presented in white size 40 Arial font on a black background. The priming task was coded using PsyToolkit (Stoet, 2017). Reaction time distributions collected via PsyToolkit and Prolific Academic

TABLE II. Interval Pairs Tested for Priming Index.

Intervals (Abbr.)	Δ Roughness	Δ Harmonicity
m3/M3	Low	Low
m6/M6	Low	Low
TT/P5	Low	High
M7/P5	Low	High
m2/P5	High	High
m2/TT	High	Low
M2/P5	High	High
m2/M3	High	Low

have been found to be comparable to RT distributions collected in a controlled laboratory environment (Armitage and Eerola, 2020).

C. Procedure

The experiment consisted of a standard word classification task with affective priming. Each item consisted of the prime (interval) presented simultaneously with a fixation cross for 250 ms. At 250 ms, the fixation cross was replaced with the target word. Participants were instructed to press the ‘z’ key if the target was negative and the ‘m’ key if the target word was positive. The target word remained onscreen for 1500 ms; key presses greater than 2000 ms after the onset of the target word were classed as timeouts. Participants initially completed a 10-item familiarisation block, which was followed by the experimental block of 32 items. During the practice block, participants were informed whether or not their response was correct immediately after each item. No indication of accuracy was given during the experimental block.

III. RESULTS

A. Data Analysis

Data pre-treatment mirrored the treatment employed in (Armitage and Eerola, 2020). RTs were fitted to an exponentially-modified (ExGaussian) distribution. RTs shorter than 250ms or slower than the 95th percentile of each participant's ExGaussian distribution were deleted prior to analysis. Similarly, incorrect answers and timeouts (i.e. RTs greater than 2000 ms) were also deleted prior to analysis. Data from participants who failed to reach a 75% accuracy rate was deleted from the analysis. Standardised effect sizes are not reported for GLM models.

B. Relationship Between Priming and Stimulus Features

Stimuli were grouped according to their acoustic properties - High vs Low difference in acoustic roughness (ΔRoughness), and High vs Low difference in harmonicity (ΔHarmonicity). Response times data were analysed using a generalised linear model. Response times in the High vs Low ΔRoughness difference conditions and the High vs Low

ΔHarmonicity conditions were compared alongside the overall congruency effect (Congruent vs Incongruent conditions). As predicted, congruency effects were present in the High ΔRoughness condition. The key interaction ΔRoughness×Congruence proved significant, $\chi^2(1) = 4.85, p = .028$. Considering the High ΔRoughness group alone, planned contrasts revealed that targets in Congruent conditions (mean RT = 667 ms, SD = 162 ms) were evaluated significantly more quickly than in Incongruent conditions (mean RT = 677 ms, SD = 162 ms), $z(inf) = 2.79$, $p = .005$, (standardised effect sizes not reported for GLM model) revealing the presence of congruency effects in this condition; contrasts between response times in the Congruent (mean RT 684 = ms, SD = 177) and Incongruent (mean RT = 682 ms, SD = 164) conditions in the Low ΔRoughness condition fell well short of statistical significance, $p > .05$. Contrary to expectations, the interaction ΔHarmonicity×Congruence was non-significant, suggesting the absence of congruency effects. Curiously, planned contrasts suggested that, in the Low ΔHarmonicity condition, there was a significant contrast in RTs between congruent (mean RT = 674 ms, SD = 161) and Incongruent (mean RT = 680 ms, SD= 157) conditions. However, when this is controlled for ΔRoughness, the contrast holds in the presence of High ΔRoughness (mean RTs = 669 vs 681 ms for Congruent vs Incongruent conditions; SD = 151 vs 151) but not Low ΔRoughness (mean RTs = 679 vs 680 ms for Congruent vs Incongruent conditions; SD = 170 vs 173), suggesting that the contrast is a consequence of the difference in roughness rather than harmonicity. Finally, the main effects of ΔRoughness, ΔHarmonicity and Congruence were all non-significant, $p > .05$. Mean response times per Congruency, ΔRoughness and ΔHarmonicity are shown in Figure 2. Within-subjects t-tests on the congruency effects for individual prime pairs are reported as supplementary material.

C. Roughness Manipulation

To probe the role of roughness further, we introduced two artificial intervals which were designed to test the influence of more extreme differences in roughness, d2 and S2 (detuned minor seconds played with piano and Shepard tone timbres). Two further sub-experiments were carried out employing the same procedure but using these artificial stimuli as primes, with the expectation that the increase in ΔRoughness would, on average, increase the difference in response times between the congruent and incongruent conditions. As the key measure of congruency effects is the difference in response times, we carried out a simple linear correlation test on difference in roughness versus difference in response time between the Congruent and Incongruent conditions (referred to from hereon as *priming index* for brevity) for the expanded data set of ten interval pairs. The correlation was statistically significant, $r = .13, t(380) = 2.53, p = .01$. However, the low r value and visual inspection of Figure 3 suggested that, rather than a linear relationship, there is a step-like relationship – i.e. the increase in ΔRoughness is *not* associated with an additional increase in the priming index. To test this, the interval pairs were split into three categories by difference in roughness: Low (m3/M3, m6/M6, M7/P5, TT/P5), High (m2/P5, m2/TT, m2/M3, M2/P5), and Extreme (made up of the manipulated intervals d2/P5 and s2/S5). Owing to the clear a priori competing hypotheses (linear vs step-like dependence on ΔRoughness), we used a planned contrasts approach, where the dependent variable was the priming index. The

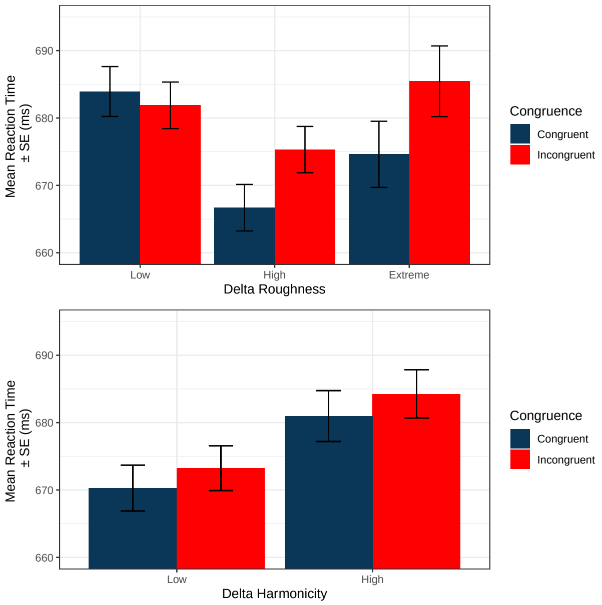

FIG. 2. Mean Reaction Times for Low, High and Extreme levels of ΔRoughness and Low and High levels of ΔHarmonicity.

planned comparisons were between the Low ΔRoughness and the combined High and Extreme ΔRoughness groups, and finally between the separate High and Extreme ΔRoughness groups.

For the first contrast, there was a statistically significant difference in the size of the priming index between the Low (mean index = -0.9ms, SD = 45.2 ms) and the combined High and Extreme groups (mean index = 10.1 ms, SD = 45.5), $t(379) = 2.47, p = .04, d = 0.24$. However, there was no significant difference in priming index between the High (mean = 9.2 ms, SD = 44.3) and Extreme (mean = 11.8 ms, SD = 48.0) ΔRoughness groups, $p > .05, d = 0.06$, suggesting that increasing the level of roughness does not increase the strength of the automatic response, supporting the hypothesis that the relationship between ΔRoughness and priming index is step-like rather than linear.

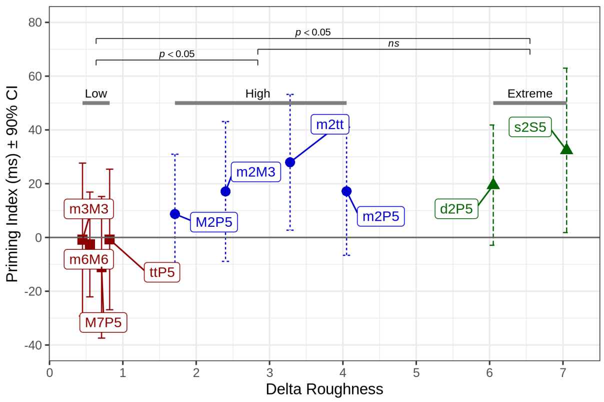

FIG. 3. Difference in Roughness vs Priming Index for Diatonic Intervals. Three groups of pairings (Low, High, and Extreme Δ Roughness) highlighting the contrasts.

IV. DISCUSSION

We have demonstrated that interval pairs which differ by only a small amount in terms of acoustic roughness do not influence responses on a word classification task. On the other hand, intervals which differ more significantly in roughness do influence the responses. Notably, out of the two acoustic components that are seen as prevalent in the perception of consonance/dissonance, namely *roughness* and *harmonicity* (see Harrison and Pearce, 2020; Parncutt and Hair, 2011), roughness was the only component that influenced the responses. Harmonicity failed to reach significance, and consequently it seems likely that harmonicity does not contribute to congruency effects. Rather surprisingly, interval pairs that bear strong cultural associations with negative and positive valence respectively (i.e. m3/M3, m6/M6, TT/P5) failed to influence responses to the classification task. This suggests that these culturally loaded interval pairs are not dissimilar enough in terms of roughness to elicit priming effects when presented by themselves, which is striking in the light that minor/major triads, in turn, have been shown to be effective primes for negative/positive words in an affective priming setting (Steinbeis and Koelsch, 2011). Also, the culturally negatively loaded interval of the tritone (see e.g. Partch, 1974) was not congruent with negative words, and was in fact an effective prime for positive words when paired with the maximally rough minor second interval due to its relatively low roughness (see Table I).

Although several contemporary models of consonance and dissonance attempt to integrate the concepts of roughness, harmonicity and other non-acoustic factors (see e.g. Harrison and Pearce, 2020; Parncutt and Hair, 2011), the present study suggests that in this context of looking exclusively at musical intervals in an affective priming setting, roughness is the most important acoustic variable. Critically, both of the intervals that are associated with increased priming index all fall within the equivalent rectangular bandwidth (see Patterson, 1976). The comparison of the minor and major thirds provides a particularly interesting case. The minor third falls theoretically just within the ERB. However, the priming index for the m3/M3 pairing did not differ significantly from zero. To probe this further, we considered whether the roughness activation was present in the basilar membrane for different combinations of partials for the two notes of each interval (see Table I). This suggests that although the fundamentals lie within the same ERB for m3, the number of overtones lying within the same ERBs is considerably less than is the case for m2 and M2, creating a much less rough effect overall. It should be noted that this approach does not account for relative amplitudes of the partials. Nevertheless, it does provide a parsimonious explanation of why an automatic response was detected for other intervals in the ERB (m2 and M2) but not for m3.

Additionally, if we consider only the 'rough' intervals, i.e. m2, M2, and the artificial intervals d2 and S2, there is no difference in the size of the congruency effects between interval pairs where the difference in roughness was high compared to extreme. This suggests that some qualitative difference exists between intervals that fall either within or outside this specific degree of ERB activation, but that beyond this threshold the degree of roughness does not influence the priming index. Indeed, the results of the present experiment provide a new behavioural correlate of the assertion by Scharf: "listeners react one way when the stimuli are wider than the critical band and another way when the stimuli are narrower" (Scharf, 1970, p. 196).

An important question is why it is roughness, and not harmonicity that influences response times. Compared to harmonicity, roughness is uniquely situated in being associated with for example alarm signals and is thought to convey an advantage in enacting automatic responses (Arnal *et al.*, 2015). A possible explanation is that it is not dissonance *per se* that is pre-activating negative concepts. The human auditory system is well attuned to human vocalisations as they carry acoustic information about for example bodily states (see Pouw *et al.*, 2020). Indeed, the harmonicity properties of the human voice have proved the basis for a compelling explanation of consonance (Bowling and Purves, 2015) if consonance and dissonance are treated as two distinct phenomena (i.e. consonance as the perceptual correlate of harmonicity, and dissonance as the perceptual correlate of roughness) instead of a continuous scale (see also Harrison and Pearce, 2020; Parncutt and Hair, 2011). As a corollary to this, we contend that the capacity to violate norms surrounding harmonicity, both in producing and attending to rough sounds, for instance cries of infants (Koutseff *et al.*,

2018) or angry voices (Bänziger *et al.*, 2015), also confers an evolutionary advantage. This advantage explains why roughness in particular and not harmonicity is associated with an increase in the priming index, i.e. a stronger behavioural response to the valenced stimuli. We argue that the automatic response present in the priming index is driven by biological adaptation to acoustic factors: acoustically rough intervals activate responses associated with acoustically rough human vocalisations such as growls, screams, or cries. Strikingly, in the case of the major second interval this seems to happen quite subconsciously, if we go by the notion of composer/theorist Paul Hindemith who proposes that the major second sounds "almost consonant to our ears" (Hindemith, 1942, p. 85). On an empirical note, the major second interval has indeed been found to be perceived as more consonant than the minor seventh and major seventh intervals (Bowling *et al.*, 2018).

Although the present study provides a novel behavioural method for tapping the roughness construct, it should be noted that at this stage the size of the effect is relatively small. Moreover, it is useful in indicating contrasts in roughness rather than as a direct measure of roughness. Nevertheless, roughness contrast is an objective new method that taps into automatic perception as opposed to aesthetic judgements and bypasses semantic confounds which have notoriously plagued consonance/dissonance research (see Lahdelma and Eerola, 2020) as well as cross-cultural research into the question (see Bowling *et al.*, 2017). Such an objective new method can help to investigate the appreciation of dissonance across musical cultures.

Indeed, the present result provides, to some extent, an explanation of the observation that, in the absence of previous exposure to Western diatonic harmony, there is no preference for consonant intervals over dissonant intervals, although there is a small preference for large over small intervals (McDermott *et al.*, 2016). The results of the present study suggest that the binary division is not whether an interval is categorised as consonant or dissonant, but whether it is large or small – i.e. falling fully within the ERB, including its partials. The present study offers an explanation of automatic responses to acoustically rough intervals based on biological imperatives to respond to alarm signals present in human vocalisation. Future research should consider whether this biological imperative underpins dissonance perception more broadly or whether this sort of priming paradigm presents a special case. The role of culture in this effect also warrants further investigation, in particular whether it can be replicated with participants who have had frequent exposure to, for instance, *beat diaphony* in musical cultures that promote roughness for its aesthetic value (Ambrazevičius, 2017; Messner, 1981).

ACKNOWLEDGMENTS

Partial financial support was provided by the Faculty Pro-Vice Chancellor's award at Anonymous University. Additional funds were provided by the Society for Education, Music and Psychology Research (SEMPRE) in the form of a Reg and Molly Buck Award awarded to the second author.

REFERENCES

Ambrazevičius, R. (**2017**). “Dissonance/roughness and tonality perception in Lithuanian traditional Schwebungsdiaphonie,” Journal of Interdisciplinary Music Studies **8**, 39–53.

Armitage, J., and Eerola, T. (**2020**). “Reaction time data in music cognition: Comparison of pilot data from lab, crowdsourced, and convenience web samples,” Frontiers in Psychology **10**, 2883.

Arnal, L. H., Flinker, A., Kleinschmidt, A., Giraud, A.-L., and Poeppel, D. (**2015**). “Human screams occupy a privileged niche in the communication soundscape,” Current Biology **25**(15), 2051–2056.

Bänziger, T., Hosoya, G., and Scherer, K. R. (**2015**). “Path models of vocal emotion communication,” PloS One **10**(9), 1–29.

Bowling, D. L., Gill, K., Choi, J. D., Prinz, J., and Purves, D. (**2010**). “Major and minor music compared to excited and subdued speech,” The Journal of the Acoustical Society of America **127**(1), 491–503.

Bowling, D. L., Hoeschele, M., Gill, K. Z., and Fitch, W. T. (**2017**). “The nature and nurture of musical consonance,” Music Perception: An Interdisciplinary Journal **35**(1), 118–121.

Bowling, D. L., and Purves, D. (**2015**). “A biological rationale for musical consonance,” Proceedings of the National Academy of Sciences **112**(36), 11155–11160.

Bowling, D. L., Purves, D., and Gill, K. Z. (**2018**). “Vocal similarity predicts the relative attraction of musical chords,” Proceedings of the National Academy of Sciences **115**(1), 216–221.

Carcagno, S., Lakhani, S., and Plack, C. J. (**2019**). “Consonance perception beyond the traditional existence region of pitch,” The Journal of the Acoustical Society of America **146**(4), 2279–2290.

Costa, M., Ricci Bitti, P. E., and Bonfiglioli, L. (**2000**). “Psychological connotations of harmonic musical intervals,” Psychology of Music **28**(1), 4–22.

Cousineau, M., McDermott, J. H., and Peretz, I. (**2012**). “The basis of musical consonance as revealed by congenital amusia,” Proceedings of the National Academy of Sciences **109**(48), 19858–19863.

Curtis, M. E., and Bharucha, J. J. (**2010**). “The minor third communicates sadness in speech, mirroring its use in music.,” Emotion **10**(3), 335–348.

Fazio, R. H., and Olson, M. A. (**2003**). “Implicit measures in social cognition research: Their meaning and use,” Annual Review of Psychology **54**(1), 297–327.

Fletcher, H., and Munson, W. A. (**1933**). “Loudness, its definition, measurement and calculation,” Bell System Technical Journal **12**(4), 377–430.

Harrison, P., and Pearce, M. (**2018**). “An energy-based generative sequence model for testing sensory theories of Western harmony,” in *Proceedings of the 19th International Society for Music Information Retrieval Conference*, Paris, France, pp. 160–167.

Harrison, P., and Pearce, M. (**2020**). “Simultaneous consonance in music perception and composition,” Psychological Review **127**(2), 216–244.

Helmholtz, H. (**1885**). “On the sensations of tone. 1954,” Trans. AJ Ellis. Introduction by H. Margenau. Republication by Dover Publ. Inc. .

Herring, D. R., White, K. R., Jabeen, L. N., Hinojos, M., Terrazas, G., Reyes, S. M., Taylor, J. H., and Crites Jr, S. L. (**2013**). “On the automatic activation of attitudes: A quarter century of evaluative priming research.,” Psychological Bulletin **139**(5), 1062.

Hindemith, P. (**1942**). *The Craft of Musical Composition, Vol. 1* (New York: Belwin-Mills).

Hutchinson, W., and Knopoff, L. (**1978**). “The acoustic component of Western consonance,” Interface **7**(1), 1–29.

Hutchinson, W., and Knopoff, L. (**1979**). “The significance of the acoustic component of consonance in Western triads,” Journal of Musicological Research **3**(1-2), 5–22.

Koutseff, A., Reby, D., Martin, O., Levrero, F., Patural, H., and Mathevon, N. (**2018**). “The acoustic space of pain: Cries as indicators of distress recovering dynamics in pre-verbal infants,” Bioacoustics **27**(4), 313–325.

Krantz, G., Merker, B., and Madison, G. (**2004**). “Subjective reactions to musical intervals assessed by rating scales,” in *Proceedings of The Eighth International Conference on Music Perception and Cognition*, Evanston, IL: Northwestern University, Vol. 5.

Lahdelma, I., Armitage, J., and Eerola, T. (**2020**). “Affective priming with musical chords is influenced by pitch numerosity,” Musicae Scientiae doi: `https://doi.org/10.1177/1029864920911127`.

Lahdelma, I., and Eerola, T. (**2020**). “Cultural familiarity and musical expertise impact the pleasantness of consonance/dissonance but not its perceived tension,” Scientific Reports (10), 8693.

McDermott, J. H., Lehr, A. J., and Oxenham, A. J. (**2010**). “Individual differences reveal the basis of consonance,” Current Biology **20**(11), 1035–1041.

McDermott, J. H., Schultz, A. F., Undurraga, E. A., and Godoy, R. A. (**2016**). “Indifference to dissonance in native Amazonians reveals cultural variation in music perception,” Nature **535**(7613), 547–550.

Messner, G. F. (**1981**). “The two-part vocal style on Baluan Island Manus province, Papua New Guinea,” Ethnomusicology **25**(3), 433–446.

Moore, B. C., and Glasberg, B. R. (**1983**). “Suggested formulae for calculating auditory-filter bandwidths and excitation patterns,” The Journal of the Acoustical Society of America **74**(3), 750–753.

Nordmark, J., and Fahlén, L. E. (**1988**). “Beat theories of musical consonance,” Q Prog Status Rep **29**, 111–122.

Oelmann, H., and Laeng, B. (**2009**). “The emotional meaning of harmonic intervals,” Cognitive Processing **10**(2), 113–131.

Parncutt, R., and Hair, G. (**2011**). “Consonance and dissonance in music theory and psychology: Disentangling dissonant dichotomies,” Journal of Interdisciplinary Music Studies **5**(2), 119–166.

Partch, H. (**1974**). *Genesis of a Music* (New York: Da Capo Press).

Patterson, R. D. (**1976**). “Auditory filter shapes derived with noise stimuli,” The Journal of the Acoustical Society of America **59**(3), 640–654.

Peretz, I. (**2010**). “Towards a neurobiology of musical emotions,” in *Handbook of music and emotion: Theory, research, applications*, edited by P. Juslin and J. Sloboda (Oxford, UK: Oxford University Press), pp. 99–126.

Pouw, W., Paxton, A., Harrison, S. J., and Dixon, J. A. (**2020**). “Acoustic information about upper limb movement in voicing,” Proceedings of the National Academy of Sciences **117**(21), 11364–11367.

Roederer, J. G. (**2008**). *The Physics and Psychophysics of Music: An Introduction* (Springer Science & Business Media).

Scharf, B. (**1970**). “Critical bands,” in *Foundations of Modern Auditory Theory*, edited by J. Tobias, **1** (Academic, New York), Chap. 5, pp. 157–202.

Schwartz, J. W., Engelberg, J. W., and Gouzoules, H. (**2019**). “Was that a scream? Listener agreement and major distinguishing acoustic features,” Journal of Nonverbal Behavior 1–20.

Smith, J. O., and Abel, J. S. (**1999**). “Bark and ERB bilinear transforms,” IEEE Transactions on Speech and Audio Processing **7**(6), 697–708.

Steinbeis, N., and Koelsch, S. (**2011**). “Affective priming effects of musical sounds on the processing of word meaning,” Journal of Cognitive Neuroscience **23**(3), 604–621.

Stoet, G. (**2017**). “Psytoolkit: A novel web-based method for running online questionnaires and reaction-time experiments,” Teaching of Psychology **44**(1), 24–31.

Terhardt, E. (**1984**). “The concept of musical consonance: A link between music and psychoacoustics,” Music Perception: An Interdisciplinary Journal **1**(3), 276–295.

Tramo, M. J., Cariani, P. A., Delgutte, B., and Braida, L. D. (**2001**). “Neurobiological foundations for the theory of harmony in Western tonal music,” Annals of the New York Academy of Sciences **930**(1), 92–116.

Vassilakis, P. N. (**2001**). “Perceptual and physical properties of amplitude fluctuation and their musical significance,” Ph.D. thesis, University of California, Los Angeles.

Wang, Y., Shen, G., Guo, H., Tang, X., and Hamade, T. (**2013**). “Roughness modelling based on human auditory perception for sound quality evaluation of vehicle interior noise,” Journal of Sound and Vibration **332**(16), 3893–3904.

Warriner, A. B., Kuperman, V., and Brysbaert, M. (**2013**). “Norms of valence, arousal, and dominance for 13,915 English lemmas,” Behavior Research Methods **45**(4), 1191–1207.