Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

PSICOLOGIA

Ciclo XXXIII

Settore Concorsuale: 11/E1

Settore Scientifico Disciplinare: M-PSI/01

MUSICAL TENSION IN HARMONIC INTERVALS: BEHAVIORAL AND NEURAL CORRELATES

Presentata da: Mattia Nese

Coordinatore Dottorato

Prof. Monica Rubini

Supervisore

Prof. Marco Costa

Co-supervisore

Prof. Andrea De Cesarei

Esame finale anno 2021

Table of contents

Abstract

Musical tension is what drives our emotional experience in music listening. However, the specific role of the musical elements involved in tension-resolution perception remains largely unclear. This dissertation aims to advance the understanding of tension perception dynamics related to sensory consonance-dissonance. Empirical works investigating the psychological and neural correlates of perceived tension will be presented.

The first experiment aimed to design and validate a new crossmodal proprioceptive device for tension rating that overcomes some of the limitations of known tools. As a result, a psychophysical equation for the matching of physical force and psychological force was presented.

The same tool was subsequently used in the second and third experiments to collect ratings of perceived tension and movement in harmonic musical intervals and standard noises. Besides, a visual analog scale (VAS) was used to allow a comparison of these two methods. The valence of stimuli was rated using only the VAS. The results confirmed the close relationship between sensory dissonance and perceived tension. Moreover, stimuli in the higher pitch register were perceived as more tense, confirming the primary role of pitch as a mediator of tension. The comparison between ratings obtained with the proprioceptive device and the VAS highlighted the tendency to give higher tension ratings using the VAS compared to the proprioceptive device. A higher perception of movement was associated with higher sensory dissonance and pitch register in harmonic intervals. High-pitch spectrum noises were associated with a higher sense of movement than low-pitch spectrum noises. Consonant intervals and low-register intervals were judged more positively than dissonant and high-pitch intervals. High-pitch spectrum purple and blue noises were evaluated as more unpleasant than low-pitch spectrum noises.

The obtained ratings were used to design the stimuli presented during the last experiment. Brain electrical activity was recorded during the presentation of short tension-resolution patterns created using the most tense (perfect unison, fourth, and fifth) and the least tense harmonic intervals (augmented fourth, minor second, and inverted major seventh) to understand how consonance-dissonance can convey meaningful information on perceived tension-resolution. Results showed overall larger effects during the 'resolution' condition compare to the 'tension induction' condition, indicating that the resolution of harmonic instability towards a state of stability may be more salient than its opposite. In particular, early-mid-latency components associated with the processing of pitch-relations between notes (P2) and context-based memory updating during the tension detection task (P3b) were found. In addition, a late positive component (LPC) was elicited, possibly reflecting deeper processing of tension-related meaning within a minimal harmonic context. While these components characterized the processing of resolution patterns, only the P3b was found during the processing of induced tension.

In sum, the present work examined the perceived musical tension resulting from shifts in sensory consonance-dissonance first by exploring its crossmodal nature to obtain subjective ratings of all harmonic intervals using a novel proprioceptive device; then highlighting the neural processing stages associated with the perception of tension-resolution patterns. Since tension is fundamental for music emotion, these studies help understand the link between music and emotion with possible insights for general theories on emotion.

General Introduction

The psychological experience of tension is a ubiquitous affective phenomenon that pervades many aspects of our lives. Music encompasses dynamic patterns of tension and resolution that characterize the emotional experience of music listening. While listening to music, it is common to feel growing expectations towards upcoming events, often associated with increased arousal. After reaching its peak, when (and if) the expectation is fulfilled, tension is usually released. This process can be repeated, creating fluctuations that run through the entire composition. Of course, the way these dynamics develop can take very different forms and influence listeners' emotional experience.

Wilhelm Wundt (1896, 1911) was the first scientist to highlight the importance of the tension-relaxation dipole (along with pleasure–displeasure and arousal–calmness) as a key factor to describe the emotional experience induced by music. A few decades later, the musicologist Leonard Meyer (1956) attempted to explain the connection between music and emotion; he presented a theory based on expectations that lead to tension. Since then, many studies have investigated tension as a key factor in understanding the link between music and emotion. However, the first general model of tension in its components was proposed by Lehne and Koelsch (2015). This model applies to the experience of music listening and many other activities and situations (e.g., reading a book, watching a movie). These authors defined tension as an affective state that (a) is associated with conflict, dissonance, instability or uncertainty, (b) creates a yearning for resolution, (c) builds on future-directed processes of expectation, anticipation, and prediction (Lehne & Koelsch, 2015). In this sense, tension cannot be assimilated to the dimension of arousal since it is possible to experience very high states of arousal without tension (e.g., winning a sports competition), since the sense of instability and uncertainty is missing, and very low states of arousal with tension (e.g., the tip of the tongue or not recalling the right name). This definition of tension is probably the most comprehensive provided by within the scientific literature. Empirical studies investigating musical tension often lacked to define it precisely or used slightly different definitions depending on their more general theoretical framework. In the next paragraph, the different definitions provided by authors of empirical studies will be discussed.

1.1 Empirical studies on musical tension

Despite being part of everyone's common experience, musical tension is an elusive concept when it comes to providing a precise and rigorous definition. It is interesting to note that empirical studies did not often provide an explicit and clear definition of musical tension to readers or participants.

From a theoretical point of view, many authors offered a description of tension based on those provided by the music theorists Meyer (1956) and Schenker (1935), who stated that ‘tension’ along with its opposite ‘relaxation’ (or ‘release’) generates ebb and flows in music listening due (at least in part) to expectancies built on present music events about future music events (Bigand et al., 1996; Krumhansl, 1996, 2002; Lehne, Rohrmeier, Gollmann, et al., 2013; Vines et al., 2006).

Regarding the instructions given to participants, many researchers voluntarily decided to not provide participants with a definition of tension during experimental procedures, relying on its self-explaining definition that may semantically recall other non-musical concepts like muscular tension or tension in social contexts (Fredrickson, 1995, 1997, 1999, 2000; Fredrickson & Johnson, 1996; Hackworth & Fredrickson, 2010; Iwanaga et al., 2005; Iwanaga & Moroki, 1999). Only in one study by Bigand and Parncutt (1999) investigating harmonic tension in chord sequences, tension was explained to participants as a feeling originating from instability that “there will be a continuation."

From a computational point of view, tension could also be defined as a result of the combination of specific music features (e.g., harmony, rhythm, timbre) thought to modulate the experience of tension in listeners. Thus, it should be possible to predict the relative amount of tension experienced during the listening of music excerpts by measuring such features and their hierarchical dependencies (Chapados & Levitin, 2008; Farbood, 2012; Farbood & Price, 2017; Lerdahl, 1996; Lerdahl & Krumhansl, 2007; Smith & Cuddy, 2003). The most comprehensive attempt in such direction was proposed by Farbood (2012), whose model includes many parameters (onset frequency, tempo, loudness, pitch height, harmony, rhythmic regularity, and meter) analyzed within discrete attentional windows.

Parallel to the above-mentioned studies, there is a line of research where tension is considered as either a discrete or a dimensional entity within models of music-evoked emotions (Eerola & Vuoskoski, 2011). For example, Zentner, Grandjean, and Scherer (2008) performed a confirmatory factor analysis resulting in a 9-factorial model of music-induced emotions, including tension along with transcendence, wonder, joyful activation, power, sadness, tenderness, nostalgia, and peacefulness. Instead, dimensional models generally consider a reduced set of broad dimensions (e.g., valence, energy arousal, tension arousal; see Schimmack & Grob, 2000), the combination of which can result in more complex nuances in the emotional experience. Eerola and Vuoskoski (2011) compared a discrete model including happiness, sadness, tenderness, fear, and anger, with a dimensional model including valence, energy, and tension. The two models did not differ in terms of the overall consistencies between the ratings. However, the discrete emotion model was less reliable when music excerpts were ambiguous examples of an emotion category compared to the dimensional model.

Another important issue characterizing empirical studies on musical tension (and more generally music-evoked emotions) concerns the distinction between *perceived* and *felt* emotion: the first refers to what we *recognize* while the latter refers to what we subjectively *experience*. Although this distinction has been often employed to describe research methods and aims, it may be less clear in reality, and it has even been suggested that the two alternatives could be seen as opposite extremes of a continuum (Gabrielsson, 2001). Recent empirical studies indicate a substantial overlapping between the two (P. Evans & Schubert, 2008; Kallinen & Ravaja, 2006; Vieillard et al., 2008).

It can be noted how musical tension has been defined in various ways depending on the specific theoretical and methodological approaches adopted (i.e., a consequence of the perception of acoustic properties, the result of expectations created by musical dependencies within a tonal hierarchy, or simply one the music-evoked emotions). This aspect may be interpreted as an intrinsic limitation for studies on musical tension, but also as a sign of its multifaceted nature and its transversality for the human experience.

Of course, the way we define the concept of tension will have important implications for the way we attempt to quantify it. Several methods were employed to allow participants to rate tension. Table 1 presents a list of empirical studies on musical tension (either perceived or felt) and the instrument employed for collecting subjective ratings. Some of these methods involve discrete retrospective ratings where participants rate the amount of tension *after* the presentation of the stimulus (Bigand et al., 1996; Bigand & Parncutt, 1999; Gingras et al., 2015; Granot & Eitan, 2011; Iwanaga et al., 1996) while others involve continuous ratings *during* music listening (Chapados & Levitin, 2008; Fredrickson, 1995, 1999; Fredrickson & Coggiola, 2003; Fredrickson & Johnson, 1996; Hackworth & Fredrickson, 2010; Lerdahl & Krumhansl, 2007; Madsen & Fredrickson, 1993; Vines et al., 2005). Some of these studies will be discussed in detail in the next chapter for their implications on the experimental studies presented here.

Table 1

Empirical Studies on Musical Tension

Study	Perceived vs. felt tension	Rating method
Sun et al. (2020)	-	Continuous rating; virtual vertical slider
Schedl et al. (2018)	Perceived	GEMS
Kleinsmith et al. (2017)	Felt	9-point Likert scale (1 = no tension/not at all; 9 = very high tension/very much.Virtual slider (left = ‘Minimal Tension’: right = ‘Maximal Tension’)
Fang et al. (2017)	Perceived	9-point Likert scale (1 = low; 9 = high)
Goodchild et al. (2016)	Perceived	Physical handheld slider; continuous rating
Gingras et al. (2016)	Perceived	Physical handheld slider; continuous rating
Choppin et al. (2016)	Felt	GEMS
Gingras et al. (2015)	Felt	7-point Likert scale (1 = very relaxed; 7 = very tense)
Sturm et al. (2015)	Perceived	Spring-loaded joystick (pushing the joystick towards = more tension; releasing the joystick = releasing tension)
Lehne, Rohrmeier, Gollmann, et al. (2013)	Felt	Continuous rating; virtual vertical slider
Farbood & Upham (2013)	Perceived	Continuous rating; virtual horizontal slider
Lehne, Rohrmeier, & Koelsch (2013)	Felt	Continuous rating; virtual slider

Farbood (2012)	Perceived	Discrete ratings; graphical shapes were depicting tension changes
Granot & Eitan (2011)	Perceived	7-point Likert scale (1 = lowest tension; 7 = highest tension)
Williams et al. (2011)	Perceived	Two-dimensional cartesian plane combining perceived tension and the amount of attention towards either melody or harmony: vertical axis from ‘Less tension’ (bottom) to ‘More Tension’; horizontal axis from ‘Melody’ (left) to ‘Harmony’ (right)
Hackworth & Fredrickson (2010)	Perceived	CRDI
Steinbeis & Koelsch (2008)	-	Signaling harmonically deviant stimuli
Margulis (2007)	Perceived	Joystick (forward = increasing tension; back = decreasing tension)
Lerdahl & Krumhansl (2007)	-	Continuous rating; virtual horizontal slider
Ilie & Thompson (2006)	Perceived	5-point Likert scale (0 = not at all tense; 4 = extremely tense)
Vines et al. (2006)	-	Continuous rating; adjustable linear slider potentiometer
Iwanaga et al. (2005)	Perceived	7-point Likert scale (1 = very little; 7 = very much)
Vines et al. (2005)	-	CRDI
Fredrickson & Coggiola (2003)	Perceived	CRDI
Smith & Cuddy (2003)	Perceived	Continuous rating; virtual vertical slider from ‘Relaxation’ (bottom) o ‘Tension’ (top)

Pressnitzer et al. (2000)	Perceived	Forced choice: tension-release, release-tension
Fredrickson (2000)	Perceived	CRDI
Iwanaga & Moroki (1999)	Felt	7-point Likert scale from 1 (very little) to 7 (very much)
Bigand & Parncutt (1999)	Perceived	11-point Likert scale from 0 (no tension) to 10 (very high tension)
Fredrickson (1999)	Perceived	CRDI
Fredrickson (1997)	Perceived	CRDI
Krumhansl (1996)	Felt	Continuous rating; virtual horizontal slider from 'Minimum' to 'Maximum'
Fredrickson & Johnson (1996)	Perceived	CRDI
Bigand et al. (1996)	Perceived	12-point scale from 1 (weak) to 12 (strong)
Fredrickson (1995)	Perceived	CRDI
Madsen & Fredrickson (1993)	Perceived	CRDI

Note. Study reference, type of tension (perceived vs. felt), and the rating method employed are included in the table. Space was left blank for studies that did not report any explicit distinction between 'perceived' and 'felt' tension. CRDI = Continuous Response Digital Interface; GEMS = Geneva Emotional Music Scale.

1.2 Event-related potentials and music processing

As the last experiment presented in this dissertation employed the event-related potential technique to investigate tension-related processes in the human brain, a short overview of the method seems appropriate.

In 1929, Hans Berger recorded the first electrical signal originating from the communication of neurons in the cerebral cortex of a human brain by placing an electrode on the scalp (Berger, 1929). Since then, the electroencephalogram (EEG) became a widely used technique for the study of brain activity with both scientific and clinical applications. However, the EEG in its original form is a combination of signals coming from many different sources and does not deliver much information as it is. In order to study the variations in brain electrical potentials that correlate in time with specific sensory, cognitive, and motor events, the raw EEG recorded during a large number of trials has to be averaged so that the signal emerges from other kinds of activation and technical noise. The brain electrical potentials extracted this way are called event-related potentials (ERPs).

In the 1960s, the study of ERP components began; scientists understood that they could investigate mind functioning by addressing broad interest research questions. Many components were found (such as the P300) that were associated with different cognitive processes. The technological progress of the last decades allowed the EEG recording and analysis to become easier, faster, and cheaper allowing an increasing number of researchers to use the ERP technique in psychological research (see Luck, 2014 for an introduction to ERP).

A classic experimental procedure to extract ERP components associated with perception and cognition involves presenting the same stimulus (or the same stimulus category) many times to one individual while recording his brain activity. Then, the time-locked signal evoked during a particular experimental condition is averaged so that the random noise is attenuated and allows the event-related component to emerge.

The human brain is involved at various stages in perception, cognition, and emotion evoked by music. In the late 1980s, the ERP technique started to be employed to explore questions about music cognition. The first study to apply ERP to the study of music processing was published by Besson and Macar (1987). This study was designed to determine whether the N400 component would be elicited by deviations involving non-linguistic expectancies. The stimuli included sentences, geometric patterns of increasing or decreasing size, scale-notes of increasing or decreasing frequency, and well-known melodies. All conditions were designed to introduce some kind of violation of expectancies. An N400 appeared only following semantic incongruities within sentences. Interestingly, the authors reported an emerging negative component peaking around 150-200 ms that we now know could have reflected processing musical expectancy violation (early right anterior negativity; ERAN; Koelsch, 2012). Similar results were later found in other experiments involving the violations of expectancies, where also P3-effects were observed (Paller et al., 1992; Verleger, 1990). Besson and Faïta (1995) found that a late positive component (LPC) was elicited when presenting musicians and nonmusicians with familiar and unfamiliar melodies with incongruous endings. The P3/LPC component seems to be related not only to the processing of violations in familiar melodies but more in general to any unexpected stimulus presented during active listening (Brattico et al., 2006).

While the studies mentioned above focused on melodies, other studies employed chords sequences to investigate the neural processing of harmonic resolutions and their violation. Thanks to these paradigms, other two components were highlighted (i.e., the ERAN and the N5) that were used to study a variety of aspects related to music-syntactic processing (Bidelman & Grall, 2014; Koelsch et al., 2000, 2007; Koelsch, 2011; Pagès-Portabella & Toro, 2020; Virtala et al., 2011, 2013; Zhou et al., 2019).

In conclusion, several commonalities among ERP studies characterizing the research on music processing can be identified: (1) the study of expectancies violation as a means to highlight the components involved at different stages of music processing (James et al., 2015; Pagès-Portabella & Toro, 2020; Steinbeis et al., 2006); (2) the comparison between musicians and nonmusicians that allows establishing what effects can be modulated by experience and what effects are independent of music learning (James et al., 2008, 2017; Jenni et al., 2017; Regnault et al., 2001); (3) experimental designs that combine language ad music to highlight broader processing mechanisms of semantic and syntactic structures (Koelsch et al., 2004, 2013; Steinbeis & Koelsch, 2008).

ERP studies whose results are of particular interest for the present dissertation will be discussed in more detail in the last chapter.

2 Perceived tension in harmonic intervals

2.1 Introduction

Musical intervals are the building blocks of musical compositions, and their distinctive acoustical properties significantly affect the musical experience. In this study, the perception of tension, movement, and valence (pleasantness/unpleasantness) in harmonic musical intervals varying in pitch-register was examined. The results are compared with the same attributes related to specific standard noises. Furthermore, the aim was to see whether the judgments of tension, movement, and valence were related to some underlying acoustical property (e.g., roughness) or exhibited distinctive and independent properties.

In the physics domain, *tension* is defined as a pulling force applied to an object. In physiology, this term mainly refers to muscle activity, namely, a state in which the muscle is contracted, as opposed to a state of muscle inactivity or relaxation. At a more metaphorical level, tension is widely used in psychology to express an emotional state of unrest, imbalance, effort, and latent hostility. Although it is often used with a negative connotation associated with fear, concern, or distress, tension could well be a property of positive emotions, such as an intense erotic desire or the expectations for an adventurous experience (Schimmack & Grob, 2000; Schimmack & Rainer, 2002).

Several studies have attempted to continuously track perceived tension in music over the whole course of a piece. The first attempt was that of Nielsen (1983), who used a pair of spring-loaded tongs with a potentiometer placed in the axis to measure the level of tension experienced during listening to Haydn's Symphony No. 104. The variations in tension during the listening task were explained in terms of grouping tendency, melodic movement, tonality, factors relating to compositional techniques, density as a function of instrumentation and sonority, dynamics as indicated in the musical score, and dynamics as assessed by a sound level meter.

Madsen and Fredrickson (1993) replicated Nielsen's research using a continuous response digital interface to record participant perception of musical tension. Unlike Nielsen (1983), the response did not require physical effort; however, the resulting tension graph showed a high degree of concordance with the one obtained by Nielsen (1983).

Several researchers employed physical and virtual sliders as tension rating instruments (Farbood, 2012; Lehne, Rohrmeier, & Koelsch, 2013; Lehne, Rohrmeier, Gollmann, et al., 2013; Vines et al., 2005). Krumhansl (1996) used a digital slider to collect participants' tension ratings while listening to Mozart's piano sonata K 282. Intersubject correlation of perceived tension was relatively high (.42), showing a good agreement among participants. Peaks of tension were recorded at the end of segments (i.e., perceived autonomous phrases within the piece). Furthermore, the highest tension peaks occurred in measures with the slowest tempos, the highest pitch in melodic contour, and relatively higher note density, dynamics, and loudness.

Lehne, Rohrmeier, Gollmann, et al. (2013) compared continuous ratings of felt musical tension for original and modified versions of two piano pieces by Mendelssohn and Mozart. Tension ratings were obtained from a virtual slider's position presented on a computer screen that could be moved with a mouse. Modifications included versions without dynamics or agogic accents and versions in which the music was reduced to its melodic, harmonic, or outer voice components. The cancellation of dynamics and agogics produced overall lower tension ratings and flatter tension profiles while preserving tension-resolution pattern. Reducing a composition to the outer voices also preserved the tension pattern, showing that the outer voices embody essential aspects of the musical structure. The authors also found a substantial redundancy between the expressive features that affected the perception of tension, which contributed to intense experiences. For example, the highest tension peaks reflected the main structural dominant on the harmonic level and were prepared by a long crescendo, the rising melody line, the lowest local bass note, the *fortissimo* and *sforzando*, and repetition of the chords.

Bigand, Parncutt, and Lerdahl (1996) investigated the effect of tonal hierarchy, sensory chordal consonance, horizontal motion, and musical training on perceived musical tension of short chord sequences. Participants had to evaluate the tension created by major or minor triads, major-minor seventh chords, and minor seventh chords when preceded and followed by a major triad on the same scale. The results showed that chords belonging to the key tonality created less tension than did non-diatonic chords. Diatonic chords falling on the first, fourth, and fifth scale degrees created a decrease in tension. The musical tension experienced on the tonic chord was weaker than that experienced on the dominant and subdominant chords. These results underline the importance of tonal hierarchies for perceived musical tension, as theorized by Lerdahl (1988, 1996) and Lerdahl and Krumhansl (2007). Bigand et al. (1996) analyzed the tonal hierarchy's role in evoking tension and found that minor and seventh chords elicited higher tension ratings than major chords. This result confirms the effect of basic acoustic features (i.e., dynamics) and timbral parameters (i.e., sensory dissonance, roughness, brightness, and density) on the perception of musical tension (Farbood & Price, 2017; Hutchinson & Knopoff, 1978; Krumhansl, 1996; Nielsen, 1983; Plomp & Levelt, 1965; Pressnitzer et al., 2000). Perceived tension tends to increase with increasing dynamics (Burnsed & Sochinski, 1998; Granot & Eitan, 2011; Ilie & Thompson, 2006; Krumhansl, 1996; Misenhelter, 2001). Among low-level timbre attributes, roughness is the most strongly related to tension. Bigand et al. (1996) reported higher roughness in tonal chord progressions correlated with higher tension. Pressnitzer et al. (2000) showed that this effect also applies to atonal harmony. Roughness is a sensation that occurs when pairs of sinusoids are close enough in frequency such that listeners experience a beating sensation. It is closely related to sensory dissonance, a term first introduced by Helmholtz (1877), who proposed that dissonance perception corresponded to the beating between partials and fundamental frequencies of two tones. Roughness is a more general term than sensory dissonance that can be applied to all kinds of sounds, including noises (Leman, 2000). Plomp and Levelt (1965) showed that roughness and sensory dissonance reached their peak when the distance between the components of a pair of pure tones was approximately one-quarter of the critical bandwidth, whose range corresponds to the 10%-20% of the center frequency (≅ three semitones) for center frequencies above 500-1000 Hz, and in the approximate range of 50-100 Hz at lower frequencies (Moore & Glasberg, 1983; Plomp & Steeneken, 1968). Later studies based on amplitude modulated (AM) tones and noises have confirmed Plomp's results, providing additional details. In these studies, the term *roughness*, rather than *sensory dissonance*, was used. Zwicker and Fastl (1990, p. 234) related roughness to three attributes: the degree of amplitude modulation, the frequency of the modulation, and the center frequency of the sound. The computation of roughness could be performed according to two main models: curve-mapping or auditory mapping. In the first case, roughness is derived from mapping all frequency intervals or frequency component pairs included in the sound spectrum. The roughness is then defined as equal to the sum of the dissonances generated by each pair of adjacent frequency components (see Sethares, 2005). The second class of models simulates cochlear mechanical filtering using an array of overlapping band-pass filters (Aures, 1985; Daniel & Weber, 1997). Vassilakis (2001) proposed a computational method that adjusted previous models balancing the relative contribution of sound pressure level and the degree of amplitude fluctuation to roughness. Hutchinson and Knopoff (1978) formalized Plomp and Levelt (1965) model so that it could be applied to musical chords. The results showed that chords with minor thirds have greater roughness than chords with major thirds, and chords with sevenths have greater roughness than chords without sevenths.

Few studies have examined the influence of pitch register on perceived tension. Granot and Eitan (2011) found that a relatively lower register (i.e., 73–139 Hz) was strongly associated with higher tension values compared to a relatively higher register (i.e., 247–466 Hz.), but only for nonmusicians. On the other hand, Ilie and Thompson (2006) found that low-pitched music was rated as less tense. However, in this study, the low (mean frequency = 156.77 Hz) and the high (mean frequency = 191.28 Hz) pitch registers were only approximately four semitones apart, whereas, in Granot and Eitan (2011), the two registers were two octaves apart. Farbood (2012) examined ascending or descending sequences of chords and found that ascending sequences were related to increased perceived tension while descending sequences were related to decreased perceived tension. However, in those sequences, the directionality of melodic pitch covaried with the register. Therefore, the evidence of the influence of pitch register on perceived tension is inconclusive. In the present studies, the pitch register was manipulated. In Experiment 2, the *low* pitch musical intervals were 19 semitones apart from the *high* set of musical intervals, while in Experiment 3, pitch varied between A0 (27.50 Hz) to C7 (2093.00 Hz). The considered ranges allowed to extend the assessment of pitch register effects on perceived tension in musical intervals.

As shown by Fredrickson (1999), having extensive familiarity with music does not substantially affect listeners' perception of tension. Both musicians and nonmusicians tend to respond similarly in tension rating tasks (Bigand & Parncutt, 1999; Fredrickson, 2000; Fredrickson & Coggiola, 2003; Frego & Frego R.J. David, 1999; Lychner, 1998), although some studies highlighted significant differences between musicians and nonmusicians. Bigand et al. (1996), for example, found that horizontal pitch motion (i.e., melodic structure) was less effective than vertical motion (i.e., changes in harmony, tonal hierarchy, and key region) in influencing the perception of tension in musicians. These results are in line with those of Parncutt (1989), who found that musicians are generally less sensitive to melodic effects and more sensitive to harmonic effects than nonmusicians. Intersubject agreement tends to be higher among musicians than nonmusicians (Bigand & Parncutt, 1999; Krumhansl, 1996). Judgments of tension also tend to be consistent for repeated trials. For example, Bigand and Parncutt (1999) noted that tension ratings were similar from the first to the fourth hearing of an excerpt.

Previous literature on musical tension perception has mainly focused on musical excerpts and chord sequences, neglecting more basic musical features such as musical intervals. Here, it is argued that an analysis of perceived tension induced by musical intervals alone could better clarify the role of sensory features such as consonance/dissonance, roughness, and brightness in comparison to more high-level musical features such as tonal hierarchy, melodic contour, or dynamic expression in the perception of tension. Musical intervals, either melodic or harmonic, are the basic units of every musical composition and profoundly impact music's expressive function (Costa et al., 2000, 2004).

In Experiment 2, musical tension ratings were expressed through two cross-modal matching procedures that were then compared. In one case, participants had to match perceived tension in musical stimuli with the muscular tension and the pulling angular movement they had to apply to a lever connected to a spring, whereas in the second case, participants had to match perceived tension with a horizontal visual analog scale (VAS). Since the pioneering work by Nielsen (1983), only two studies have employed a proprioceptive system based on force feedback for the evaluation of auditory sensations (e.g., loudness) (Susini et al., 2002; Susini & McAdams, 2000). It is argued that the mapping of perceived tension over muscular tension, using a lever with a wide rotation angle (60°), would lead to more accurate ratings since the two dimensions (i.e., perceived tension and muscular tension) shared the same core concept of tension. Furthermore, a cross-modal matching procedure has the advantage of avoiding biases associated with numerical ratings (Link et al., 1991) and the effects of pitch mapping on the horizontal and vertical space. The rating could be affected by the sensory mapping of pitch on the vertical and horizontal space dimensions with judgments expressed in a vertical or horizontal direction. Pitch has a primary space mapping on the vertical space (Bonetti & Costa, 2018; K. K. Evans & Treisman, 2011; Rusconi et al., 2006), and secondary mapping on the horizontal dimension (high pitch-right, low pitch-left) in musicians (Rusconi et al., 2006).

The three studies presented in this chapter aimed to investigate how perceived tension, perceived movement, and pleasantness varied across musical intervals and standard noises and how it was modulated by pitch register using two cross-modal matching psychophysical procedures. In Experiment 1, a psychophysical calibration of a proprioceptive device used for the subsequent perceived tension ratings was performed. The power function relating the physical force to the apparent force was determined using a ratio production task (S. S. Stevens, 1959; Susini & McAdams, 2000) in which participants had to double or halve a given initial tension that varied for each trial. In Experiment 2 and 3, two cross-modal matching procedures were used to assess perceived tension, perceived movement, and pleasantness in musical (harmonic) intervals and five standard noises. All musical intervals within the octave were considered, including the unison (13 intervals). Also, brown, pink, white, blue, and purple noise were evaluated. The choice to include standard noises was the opportunity to test on a broader psychoacoustic scale the role of roughness and brightness in the perception of tension, movement, and valence. If the perception of tension is mainly due to roughness, standard noises should be perceived as extremely tense, as they consist of many frequencies falling within critical bandwidths, which give rise to the mixing of many beating sounds. The pitch register was also manipulated, comparing a set of intervals in a low-pitch register with a set of analogous intervals transposed to a high-pitch register. The timbre was a combination of fundamental and five harmonics with a linear decreasing amplitude. In Experiment 2, participants had to assess the perception of tension and movement with the proprioceptive device, whereas valence (pleasantness-unpleasantness) was rated with a VAS. In this case, the VAS was chosen so that a neutral-central point could be set, whereas, in the proprioceptive device, the scale must be unipolar with a null point and a maximum point. Experiment 3 mirrored the experimental procedure used in Experiment 2 except for the cross-modal assessing device since a VAS was used for all the three dependent variables: perceived tension, perceived movement, and valence.

2.2 Experiment 1: preliminary validation of a new crossmodal proprioceptive device for the rating of tension

In order to design a cross-modal matching task between perceived tension, perceived movement in musical stimuli, and muscular tension/angular movement, a proprioceptive device was developed, consisting of a long lever (85 cm) with an angular displacement of 60° and a pulling range of 0 - 33.1 N, as showed in Figure 1. A relatively long lever with a high angular excursion was employed to maximize the rating range. The first experiment aimed to provide a psychophysical validation of the proprioceptive device, determining the power function linking the physical force (expressed in Newtons) to the apparent force using a ratio production task (S. S. Stevens, 1959; Susini & McAdams, 2000). The procedure mirrored the one used by Susini and McAdams (2000) for the rating of loudness.

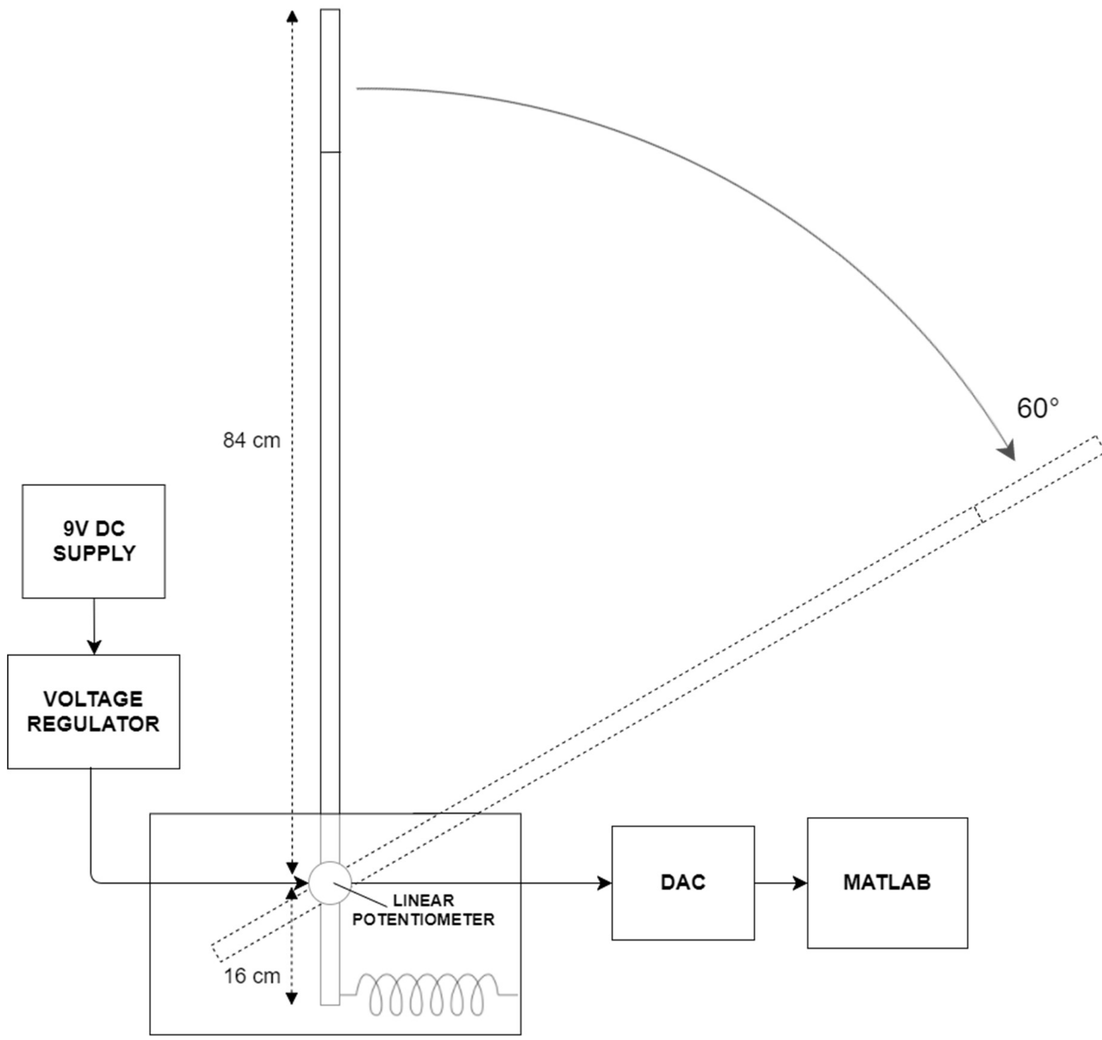
2.2.1 Method

Participants

Eighteen university students participated in the experiment (8 females, M_{age} = 24.78 years, *SD* = 6.82). All participants were right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). The students participated voluntarily.

Figure 1

The Proprioceptive Device Developed for Perceived Tension and Movement Ratings



Note. The lever's angular displacement from the initial position is accompanied by the increasing force exerted by the participant’s hand. A linear 10 kΩ potentiometer mounted on the rotation fulcrum of the lever modulated a 9V DC supply. A DAC converted the output voltage, and the data recorded with a Matlab script.

Apparatus

The proprioceptive device (Figure 1) consisted of a vertical lever (85 cm) rotating around a fulcrum. A harmonic-steel spring was attached (16-cm apart from the fulcrum) to the lever's lower end and horizontally joined to the metallic chassis. When the upper end of the lever was pulled, the spring created a linearly increasing tension. The maximum rotation displacement of the lever was 60°, corresponding to a force of 33.1 N. In order to continuously record the displacement of the lever, a 9V DC supply and a voltage regulator (to maintain a constant voltage) were wired to a linear 10 kΩ potentiometer mounted in the fulcrum so that the output voltage was a linear function of the lever displacement. Voltages were converted into digital values using a DAC device (National Instruments USB-6225) and recorded on a PC using a Matlab script. Red tape was applied to the top 10 cm of the lever, marking the participants' handgrip position.

The linearity of the function relating the output voltage with the physical force was computed sampling the physical force (*N*) and the output voltage (*V*) over 20 discrete angles equidistant from each other (3°), covering the whole 60° displacement. The force was measured with a digital dynamometer (accuracy: ± 0.049 *N*). The resulting linear regression had an R^2 of .993. The linear function is reported in Equation 1, where *N* is for force expressed in Newtons, and V is the voltage measured at the potentiometer output.

$$N = 0.331V \tag{1}$$

Procedure

Written informed consent was obtained from all participants before the beginning of the experiment. Participants were seated comfortably on a chair in front of a computer with the lever on their right. After explaining how the proprioceptive device worked, participants were asked to familiarize themselves with the lever. They were instructed about the ratio production task and asked to follow the screen's instructions during the experimental session. Each session included two conditions (25 trials each, including five practice trials) for 50 trials in total. At the beginning of each trial, participants were required to pull the lever until they heard a continuous beeping sound. Then, starting from that position (position A), they were required to double the tension ('double' condition) or to halve the tension ('halve' condition). Once they reached the target position (position B), they had to press the spacebar on a keyboard and move the lever to the initial rest position waiting for the following trial. A five-second pause after each trial ensured muscular rest. The order of the two conditions ('double' and 'halve') was counterbalanced across participants. Position A was randomly assigned in each trial (tolerance ± 1.2°) using a specific restriction in the 'double' condition: position A randomly varied between 0° and 27° (corresponding to the 45% of the overall angle of displacement) to avoid a ceiling effect.

Data Analysis

Voltage values recorded during the ratio production task were converted into force values (expressed in *N*) using Equation 1. The constant *k* and the exponent *a* were estimated using the Curve Estimation function in SPSS. The curve estimation was performed separately for the 'double' and 'halve' conditions. The force values (*N*) that matched positions A were doubled in the 'double' condition and halved in the 'halve' condition. Then these values were regressed with the apparent forces that matched positions B. A general power function was then obtained by averaging the individual values the constant *k* and the exponent *a* for the two conditions.

2.2.3 Results

The mean exponent of the power function was 1.03, while the mean constant was 1.18. Thus, the resulting proprioceptive power equation is reported in Equation 2, where Ψ designates the apparent tension and Φ the physical force. Since the exponent is greater than

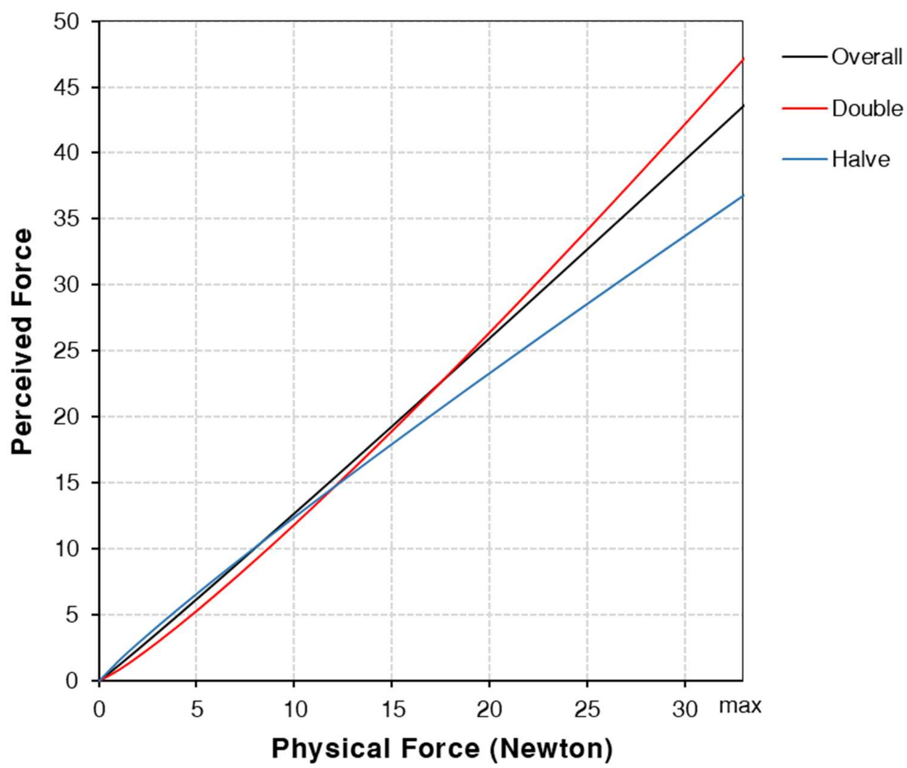
1.00, the proprioceptive tension sensation was positively accelerated as a function of the magnitude of physical force.

$$\Psi = 1.18(\Phi^{1.03}) \tag{2}$$

Figure 2 shows the power functions of the two experimental conditions ('double' and 'halve') and the general power function resulting from the average of the two conditions.

Figure 2

Power Function Relating the Physical Force to the Perceived Force



Note. The black line represents the overall power function: $y = 1.18(\mathrm{x}^{1.03})$. The red line represents the power function obtained in the 'double' condition: $y = 0.83(x^{1.16})$; while the blue line represents the power function obtained in the 'halve' condition: $y = 1.53(x^{0.91})$.

2.2.4 Discussion

In this experiment, a psychophysical calibration of a novel cross-modal proprioceptive device for tension ratings in musical stimuli was presented. The calibration followed a ratio production procedure in which participants had to double or halve a given pulling force applied to the lever (S. S. Stevens, 1958). The power function relating to physical and perceived force showed a general tendency to overestimate the force applied to the proprioceptive device, with a greater overestimation in the 'double' condition than the 'halve' condition. This effect increased as the physical force increased.

Previous studies investigating the power function for muscular tension are not unanimous in showing a specific exponent. J. C. Stevens (1989), for example, presented several psychophysical functions of isometric force with exponents ranging between 1.5 and 1.8 . S. S. Stevens (1975) measured the apparent muscular force exerted by a participant on the handle of a dynamometer using different judgment methods obtaining a power law with an exponent of 1.7 . In another study, a force was applied to the palm, yielding an exponent of 1.1 (S. S. Stevens, 1960). Susini and McAdams (2000) validated a proprioceptive device in which both the force and the angular displacement varied, obtaining an exponent of 1.77. On the contrary, Van Doren (1996) found exponents between 0.6 and 0.8 in a halving and doubling procedure for the isometric force assessment. The differences highlighted in previous literature could be attributed to the high variability in the methods and procedures for eliciting muscular force and differences in the scaling techniques (Link et al., 1991).

The same device described in this experiment was used to evaluate perceived tension and perceived movement of musical stimuli in Experiment 2.

2.3 Experiment 2 and 3 - Subjective ratings of tension, valence, and movement in harmonic intervals and noises

In experiments 2 and 3, we applied two cross-modal procedures for studying the perception of tension, movement, and pleasantness/unpleasantness of harmonic musical intervals and standard noises. The device described in Experiment 1 was used to assess the perceived tension and movement of musical stimuli in Experiment 2, whereas, in Experiment 3, the same stimuli were evaluated using a VAS. In both experiments, stimuli were musical intervals (all the musical intervals from the unison to the octave) and five calibrated noises (white, purple, blue, pink, and brown) differing in their spectrum and emphasis on low-pitch or high-pitch frequencies. Two different pitch registers (*high-pitch* and *low-pitch*) were used to create two sets of intervals. Pitch register was introduced as an independent variable as only a few studies considered its effect on perceived tension (Farbood, 2012; Granot & Eitan, 2011; Ilie & Thompson, 2006). Perceived tension, movement, and valence were also compared with the level of roughness of musical intervals and noises, computed according to Sethares (2005).

2.3.1 Method

Participants

Experiment 2: Twenty-five university students (17 females, M_{age} = 25.47 years, *SD* = 6.82, and 8 males: M_{age} = 24.50 years, *SD* = 1.87) participated in the experiment. None of the participants was a professional musician. The distribution of years of music experiment or musical instrument practice between participants was: 0 years: 18, 1 year: 2; 3 years: 2, 5 years: 1, 6 years: 1.

Experiment 3: Twenty university students (5 females, M_{age} = 29.83 years, *SD* = 10.03, and 15 males: M_{age} = 27.07 years, *SD* = 7.89) participated in the experiment. The distribution of years of music study or musical instrument practice between participants was: 0 years: 12; 1 year: 1; 2 years: 4; 3 years: 1; 5 years: 1; 10 years: 1.

For both experiments, none of the participants had hearing loss (self-reported). Participation was on a voluntary basis and informed written consent was obtained from each participant. Both experiments were approved by the University of Bologna research ethics committee.

A statistical power analysis was performed for sample size estimation. With an alpha = .05 and power = .95, the projected total sample size needed with an effect size = 0.15 (GPower 3.1) is approximately N = 38 for a within-between group comparison. Thus, our proposed sample size of N = 45 will be more than adequate for the main objective of this study.

Apparatus

The proprioceptive device validated in Experiment 1 was used for tension and movement ratings in Experiment 2. The audio output was controlled by a USB Audio/MIDI interface (Roland UA-25). Audio stimuli were delivered over noise-isolating headphones (Sennheiser HD 2.20s). Stimuli were presented using the E-Prime software. Tension and movement ratings were acquired and recorded through a Matlab script. Synchronization between the E-Prime software and the Matlab acquisition routine was guaranteed by a parallel-port connection between the two PC. Valence ratings were acquired through a horizontal VAS from 'Unpleasant' (left) 'Pleasant' (right), presented at the center of the screen (viewing angle of 11.6°) with a cursor that could be moved with the mouse. The presentation of the stimuli and the collection of the responses were controlled through a computer running Matlab with PsychToolbox 3 (Brainard, 1997). In Experiment 3, the VAS was used for all the ratings (tension, movement, and valence).

Stimuli

Two sets of thirteen musical dyads (i.e., two-note musical intervals) were digitally created using the Csound software. One set comprised the thirteen musical intervals within an octave (from the perfect unison to the perfect octave) using C3 as root note (*low-pitch* condition); the other set was analogously built using G4 as root note (*high-pitch* condition). The characteristics of harmonic intervals employed in the current study are presented in Table 2.

Table 2

Features of Harmonic Intervals on C3 and G4 Root

N° of semitones	Name	Pitch ratio	Upper note fundamental frequency (C_3)	Upper note fundamental frequency (G_4)
0	Perfect unison (P0)	1:1	130.81	392.00
1	Minor second (m2)	16:15	139.53	418.13
2	Major second (M2)	9:8	147.16	441.00
3	Minor third (m3)	6:5	156.97	470.40
4	Major third (M3)	5:4	163.51	490.00
5	Perfect fourth (P4)	4:3	174.41	522.67
6	Augmented fourth (A4)	45:32	183.95	551.25
7	Perfect fifth (P5)	3:2	196.22	588.00
8	Minor sixth (m6)	8:5	209.30	627.20
9	Major sixth (M6)	5:3	218.02	653.33
10	Minor seventh (m7)	16:9	232.55	696.89
11	Major seventh (M7)	15:8	245.27	735.00
12	Perfect octave (P8)	2:1	261.62	784.00

Note. For each interval, the number of semitones between the two notes, the name of the interval, the pitch ratio, and the fundamental frequency of the upper note in the two registers ("low" and "high"), are reported.

Stimuli in the two conditions were, therefore, 19 semitones apart. Intervals were computed using just ratios between the lower and the upper voice (five-limit tuning) to exclude the presence of beatings due to a specific tempered tuning system. The frequency spectrum of each note forming the dyads was computed adding five linear-decreasing partials to the fundamental frequency according to Formula 3 (Bidelman & Krishnan, 2009; Plomp & Levelt, 1965),

$$f_0 + \frac{1}{2}f_1 + \frac{1}{3}f_2 + \frac{1}{4}f_3 + \frac{1}{5}f_4 + \frac{1}{6}f_5 \tag{3}$$

An example of the frequency spectrum resulting from a perfect fifth interval is shown in Figure 3. The standard noises include white, purple, blue, pink, and brown noises. In white noise, all 20-20000 Hz frequencies had equal power. Purple noise power density increased 6 dB per octave with increasing frequency (density proportional to f^2). Blue noise power density increased 3 dB per octave with increasing frequency (density proportional to f). In purple and blue noises, high-register frequencies were dominant. In pink noise, there was a fall off of 3 dB/octave in power density with increasing frequency (density proportional to 1/f). The frequency spectrum was linear on a logarithmic scale. In brown noise (also Brownian or red noise), the power density decreased 6 dB/octave with increasing frequency (density proportional to $1/f^2$). In pink and brown noises, low-register frequencies were dominant. The spectrums of the five noises, considering a linear frequency scale in abscissa, are reported in Figure 4. Amplitude is reported as relative magnitude among frequencies, as spectra were produced from digital sound files.

Stimuli were stationary sounds with a rise- and decay-time of 50 ms; their loudness was equalized to 23.88 sones with the Matlab Genesis Loudness Toolbox (Genesis, 2009), applying the ANSI S34 2007 procedure (American National Standards Institute, 2007). Stimuli were presented at a sound level of 68.5 dBA (measured with a DeltaOhm HD2010 phonometer set with A ponderation curve). All the stimuli (musical intervals and noises) are available in the Supplementary material. According to Sethares' model, roughness values were computed using the MIRtoolbox for Matlab (Lartillot et al., 2008). The levels of roughness for musical intervals and standard noises used in Experiments 2 and 3 are shown in Figure 5. For the informative purpose, the brightness level assessed through MIRtoolbox 1.7.2 (Lartillot et al., 2008) is shown. Brightness is related to the amount of energy that exceeds a specific frequency threshold that, in this case, was set as 1500 Hz. It was expressed as a proportion ranging from 0 to 1 (Figure 6). The correlation between roughness and brightness was .47 ($p < .001$). The mean roughness level for noises was higher than for musical intervals (M_{noise} = 5952.99, $M_{intervals}$ = 725.983). The difference was significant: $F(1, 29) = 10.08$, $p = .003$, $\eta_p^2 = 0.26$.

Considering musical intervals only, roughness was higher for low-pitch intervals (M = 1352.73) than for high-pitch intervals (M = 99.23): $F(1, 24) = 28.96$, $p < .001$, $\eta_p^2 = 0.54$. Roughness parameters were not significantly different between consonant intervals (M = 904.99), imperfect consonant intervals, thirds and sixths (M = 1006.93) and dissonant intervals (M = 763.64): $p = .68$.

Figure 3

Frequency Spectrum of the Perfect Fifth Interval on C3 Root

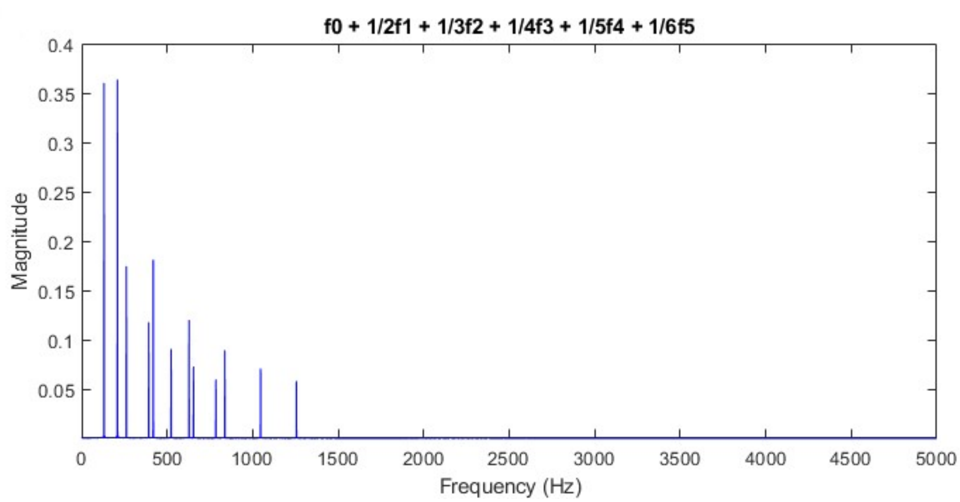


Figure 4

Frequency Spectrum of Brown, Pink, White, Blue, and Purple Noises

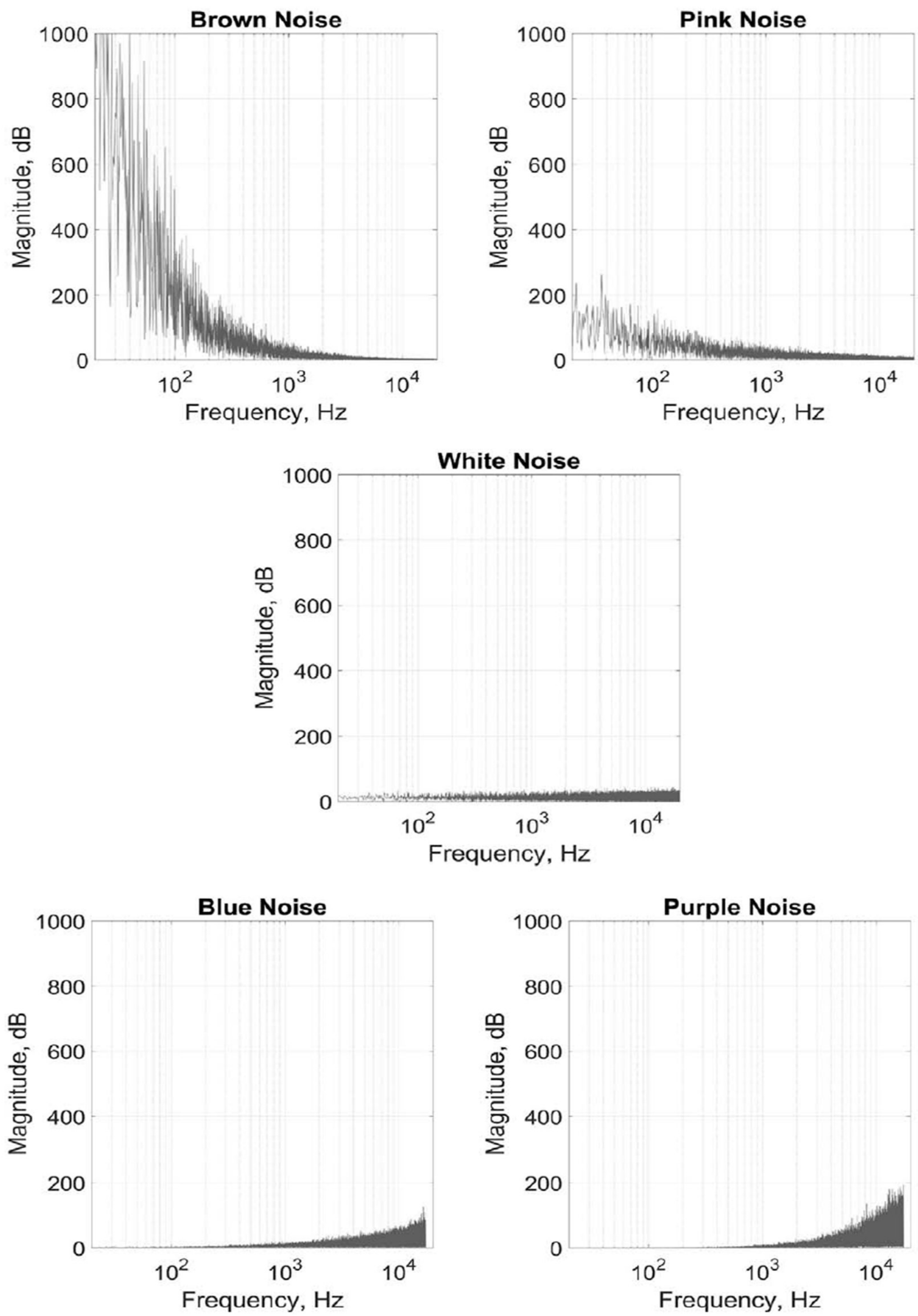


Figure 1

Roughness Level of Intervals and Noises Employed in Experiment 2 and 3

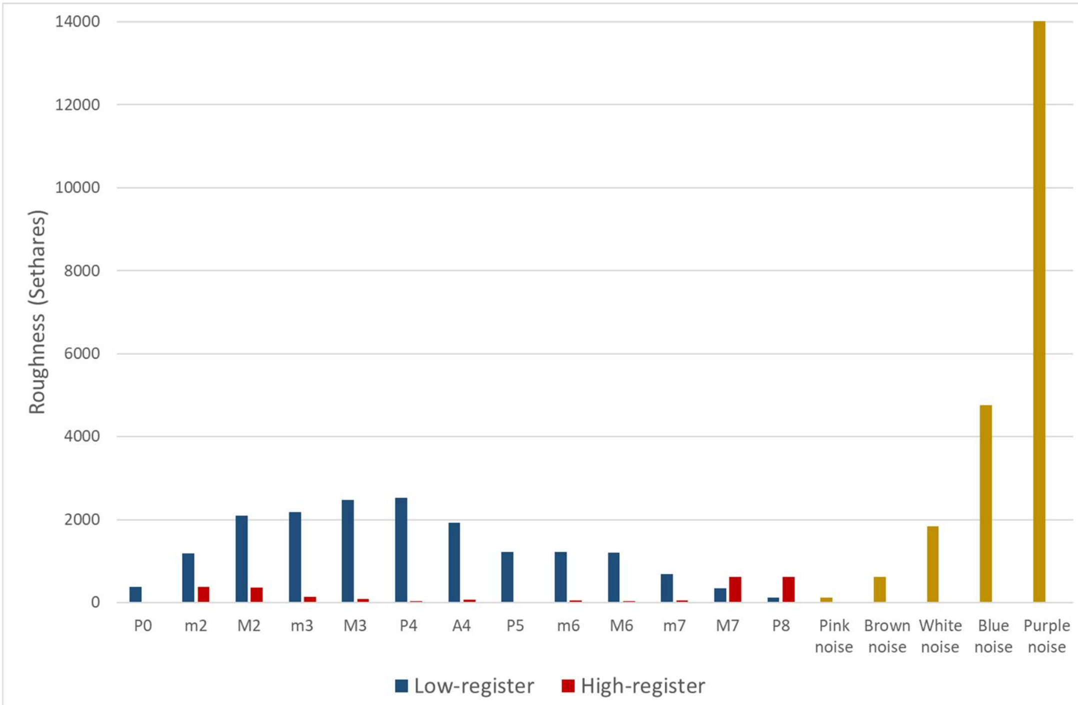
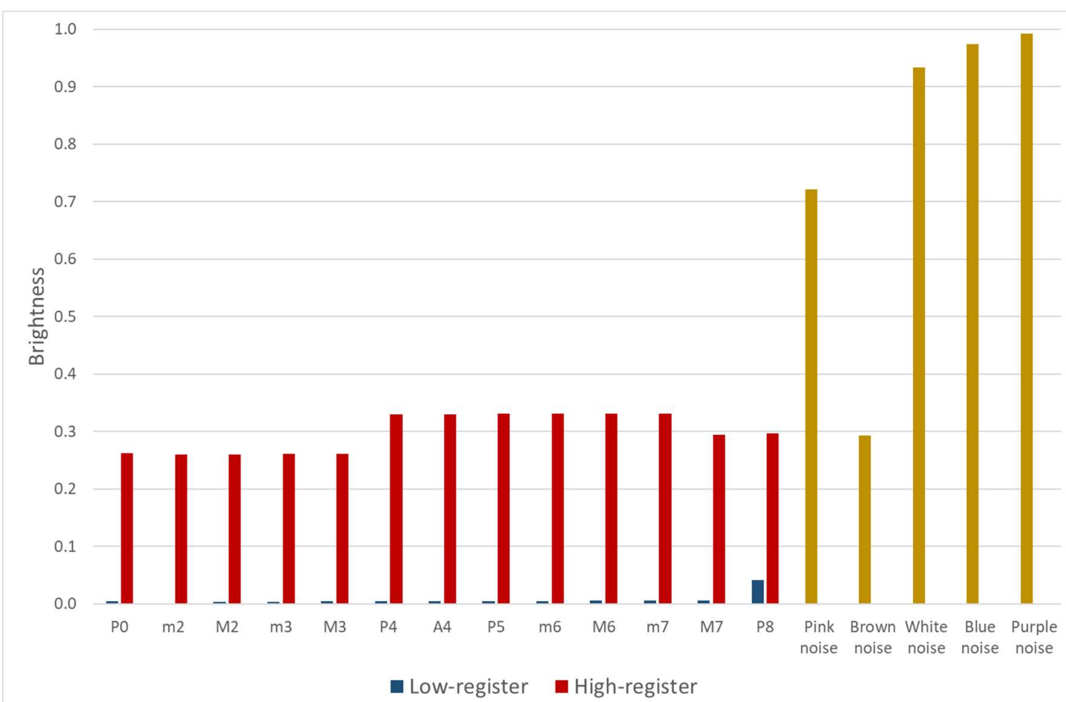


Figure 2

Brightness Level of Intervals and Noises Employed in Experiment 2 and 3



Procedure

A brief questionnaire assessed self-reported hearing problems and the years of music practice (singing or playing an instrument) and music study. Ratings for tension, pleasantness, movements were collected in three separate blocks, whose order was randomized between participants. Each block consisted of 36 trials in which the 13 musical intervals and the five standard noises were presented twice. The order of trials was randomized assigned within each block, and the order of blocks (tension, movement, valence) was randomized for each participant. Each stimulus had a duration of 1 second and could be relistened by pressing the 'R' key on the keyboard, although participants were encouraged to not relisten each sound many times. As stated before, the proprioceptive device described in Experiment 1 was used for the assessment of the perceived tension and movement of musical stimuli in Experiment 2, whereas in Experiment 3 the same stimuli were evaluated using a visual analog scale. Valence was rated by means of a visual analog scale in both experiments as its bidimensional nature could be poorly matched by the force applied to the proprioceptive device. Participants gave their response by There was an inter-stimulus-interval of 3.5 s. The response was not time-limited.

Data analysis

The ratings produced with the proprioceptive device were obtained by converting the output voltages into pulling force (expressed in Newtons) using Equation 1 reported in Experiment 1. The pulling force was then converted to perceived force with Function 2 found in Experiment 1. The maximum perceived force, corresponding to a lever pulling until the upper limit, was 43.65. For the VAS ratings, the chosen point was converted as a percentage of line bisection. Therefore, ratings were transformed into values between 0 (left) and 100 (right).

The data were analyzed considering these independent variables: (a) crossmodal procedure with two levels (proprioceptive device and VAS); (b) stimulus with 18 levels (13 musical intervals and five noises); (c) pitch register with two levels (low-pitch and high-pitch, only for musical intervals); (d) consonance with three levels (perfect consonances: P0, P4, P5, P8; imperfect consonances: m3, M3, m6, M6; dissonances: m2, M2, A4, m7, M7). Dependent variables were the ratings of tension, movement, and valence. Roughness and years of musical studies/instrumental practice were included in the model as covariates. Pairwise comparisons were performed using the Tukey-HSD test.

Ratings collected with the proprioceptive device were transformed from a 0-43.65 range to a 0-100 range to directly compare them with the VAS ratings. Figures 7, 8, and 9 show the distributions for perceived tension, movement, and valence ratings for both Experiments 2 and 3. The distribution of valence ratings presented a relatively low skewness (-.01 and .005 for Experiment 2 and Experiment 3, respectively), and the levels of kurtosis were negative (-.19 and -.55 for Experiment 2 and Experiment 3) (Figure 7). The distribution of perceived tension ratings obtained with the proprioceptive device was positively skewed (.42), while the distribution related to the use of the VAS was negatively skewed (-.22). At the same time, kurtosis was negative in both cases (-.48 and -.83 when using the proprioceptive device and the VAS, respectively) (Figure 8). The distributions of perceived movement ratings were positively skewed for both rating methods (.65 when using the proprioceptive device and .18 when using the VAS); the distribution presented a negative kurtosis when using the VAS (-.83) and a positive kurtosis (.66) when using the proprioceptive device (Figure 9).

The data were analyzed by applying a linear mixed-effect model (Laird & Ware, 1982; Pinheiro & Bates, 2000). The assumption of normality of residuals was tested with a visual inspection of the Q-Q plot. For each dependent variable (valence, tension, and movement), two linear mixed model analyses were performed: the first including all stimuli (musical intervals and noises) and the second including only musical intervals because attributes (i.e., high-pitch vs. low-pitch register and level of consonance) pertained only to musical intervals and not to noises. In both analyses, participants were considered as a random effect.

In the analysis involving all stimuli, the type of stimulus and the crossmodal procedure (VAS vs. proprioceptive device) were entered as fixed effects, while roughness and years of musical study/instrumental practice were included as covariates. In the analysis involving musical intervals, the degree of consonance (perfect consonance, imperfect consonance, dissonance), the register (high, low), and the crossmodal procedure were entered as fixed effects, whereas roughness was entered as a covariate. Each fixed effect was included sequentially in the model to test if it contributed significantly or not to increase the model's validity. Each model was fit by maximizing the log-likelihood and assessed using the Akaike information criterion (AIC). Valence ratings are pictured on a scale from -50 to +50 to increase the results' legibility since the scale's midpoint indicated neutral valence. All statistical computations were performed using *R* (version 3.6.1).

Figure 7

Distribution of Valence Ratings in Experiments 2 and 3 (VAS)

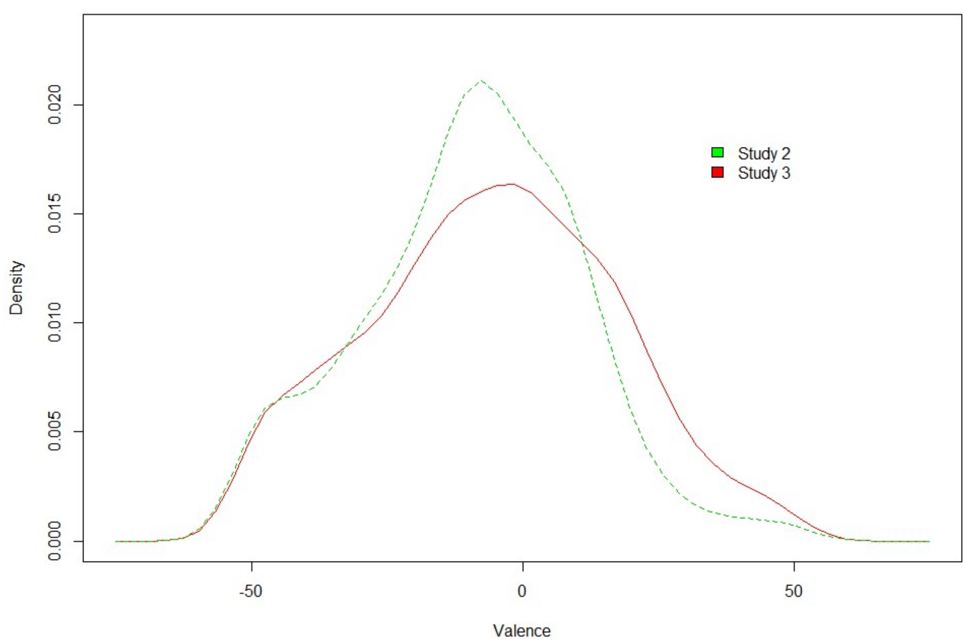


Figure 8

Distributions of Perceived Tension Ratings in Experiments 2 and 3

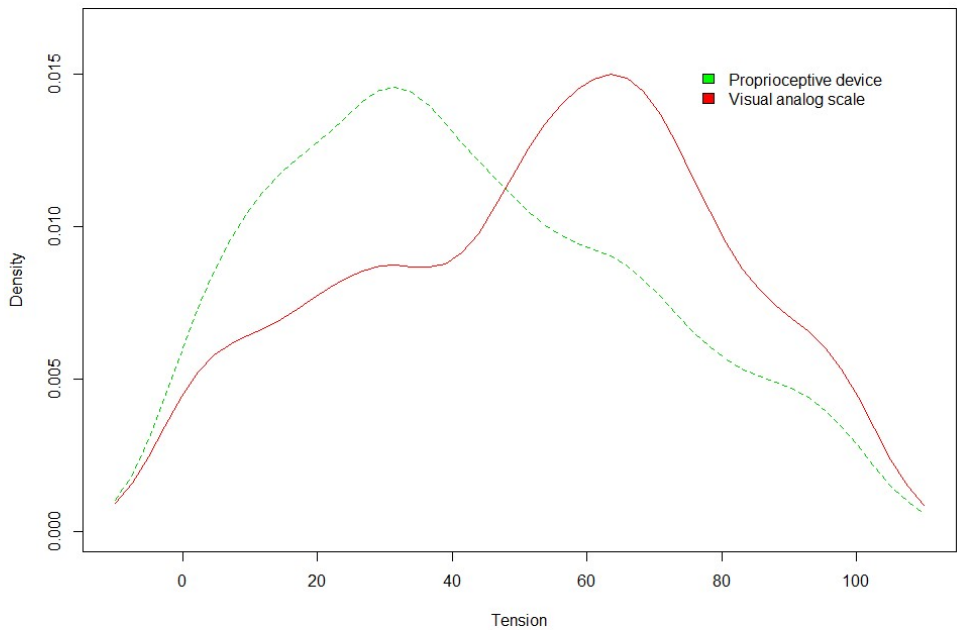
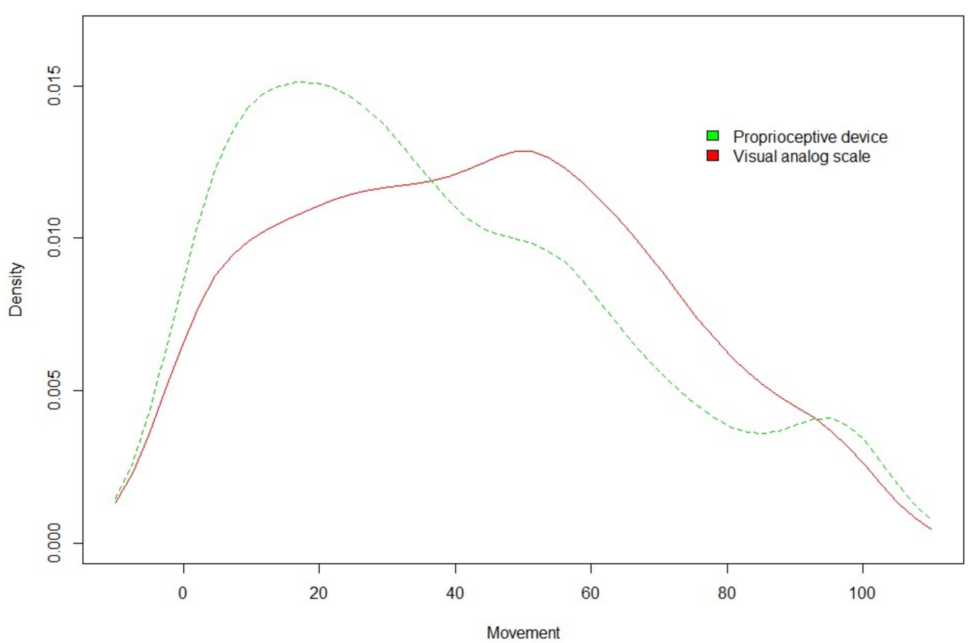


Figure 9

Distributions of Perceived Movement Ratings in Experiments 2 and 3



2.3.2 Results

Correlations

Pearson's correlations between the dependent variables and the covariate roughness are reported in Table 3. Valence was negatively correlated with tension and roughness. The correlation with movement was very low but significant. Tension was negatively correlated with valence, and positively correlated with movement. The correlation between tension and roughness was not significant. Movement was slightly negatively correlated with valence, and positively correlated with tension and roughness.

Valence – Musical intervals and noises

Table 4 shows the comparison between the incremental linear mixed models that tested the fixed effects on valence ratings. The stimulus and the covariate roughness showed significant effects on valence ratings. Table 5 shows all the parameters and coefficients included in the model. The Q-Q plot of the residuals is shown in Figure 10. Estimated marginal means and 95% confidence intervals for valence ratings considering the stimuli used in Experiments 2 and 3 are reported in Figure 11.

Valence – Musical intervals only

The linear mixed model testing the effects of consonance, register, roughness, and crossmodal procedure on valence ratings of musical intervals showed a significant effect of consonance and pitch register, as reported in Table 6, which shows the estimated parameters of the model. Estimated means and 95% confidence intervals for the three levels of consonance as a function of pitch register are reported in Figure 12.

Estimated marginal means for the three consonance levels were: consonant intervals: -2.43 (*SE*: 1.62); imperfect consonance: -5.79 (*SE*: 1.62); dissonance: -11.41 (*SE:* 1.59). Tukey HSD tests showed that all the contrasts between the three levels were significant. High-register intervals were evaluated as more unpleasant (*EMM:* -11.81; *SE:* 1.61) than low-register intervals (*EMM:* -1.28; *SE:* 1.61) ($z = 10.92$, $p < .001$).

Table 3

Spearman Correlations Between the Dependent Variables and Covariates

	Tension	Movement	Roughness
Valence	-.28 ***	-.06*	-.13**
Tension		.19**	-.03
Movement			.08**

**p* < .05. ***p* < .01. ****p* < .001

Table 4

Linear Mixed Model Results for the Valence Ratings

Model	*df*	*AIC*	χ^2	*p*
1. Intercept	3	24136.73		
2. 1 + Roughness	4	24076.47	62.25	< .001
3. 2 + Years of musical study/practice	5	24075.78	2.68	.10
4. 3 + Crossmodal procedure	6	24076.02	1.76	.18
5. 4 + Stimulus	23	23746.23	363.79	< .001

Note. The fixed factors were sequentially included in the model. The participant was considered a random factor.

Table 5

Parameter Estimates in the Linear Mixed Model for Valence Ratings Including Roughness and Stimulus as Predictors

Predictor	*B*	*SE*	*t*	*p*
(Intercept)	-4.27	1.87	-2.28	.02
Roughness	-0.05	0.007	9.87	< .001
m2	-20.26	1.97	-10.24	< .001
M2	-13.51	2.02	-6.66	< .001
m3	-13.09	2.01	-6.48	< .001
M3	-7.01	2.03	-3.45	< .001
P4	-5.94	2.03	-2.92	.003
A4	-11.00	1.99	-5.50	< .001
P5	0.61	1.96	0.08	.93
m6	-5.33	1.96	-2.71	.006
M6	-0.95	1.96	-0.48	.62
m7	-4.02	1.95	-2.05	.03
M7	-7.76	1.95	-3.96	< .001
P8	0.32	1.95	0.16	.86
Pink	2.24	2.39	0.93	.34
Brown	5.16	2.40	2.15	.03
White	-17.69	2.54	-6.96	< .001
Blue	-40.03	3.36	-11.90	< .001
Purple	-95.25	7.80	-12.20	< .001

Figure 10

Q-Q Plot for the Linear Mixed Model Referred to Valence Ratings

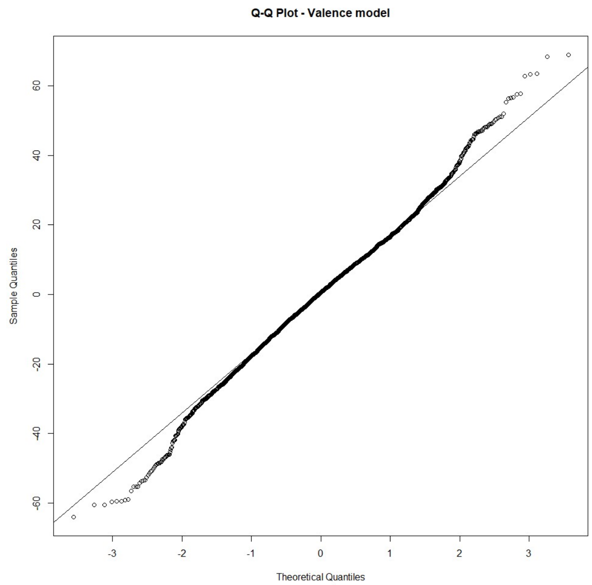


Figure 11

Estimated Marginal Means and 95% Confidence Intervals for Valence Rating of Each Musical Interval and Noise

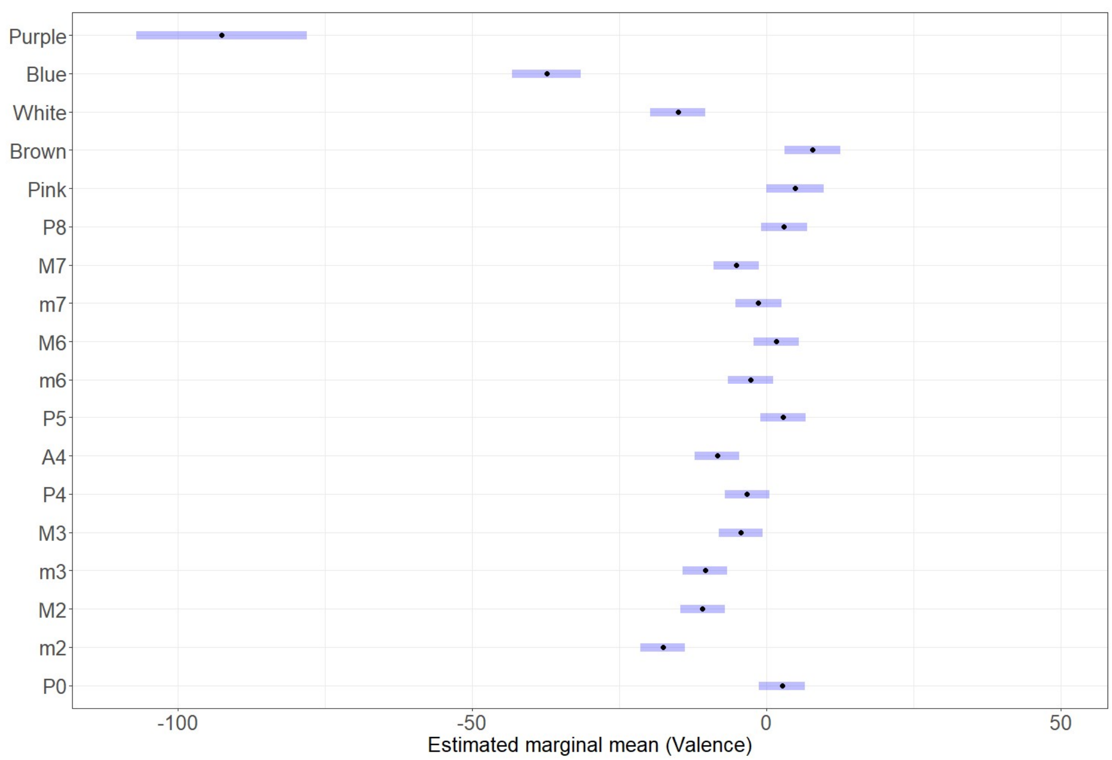


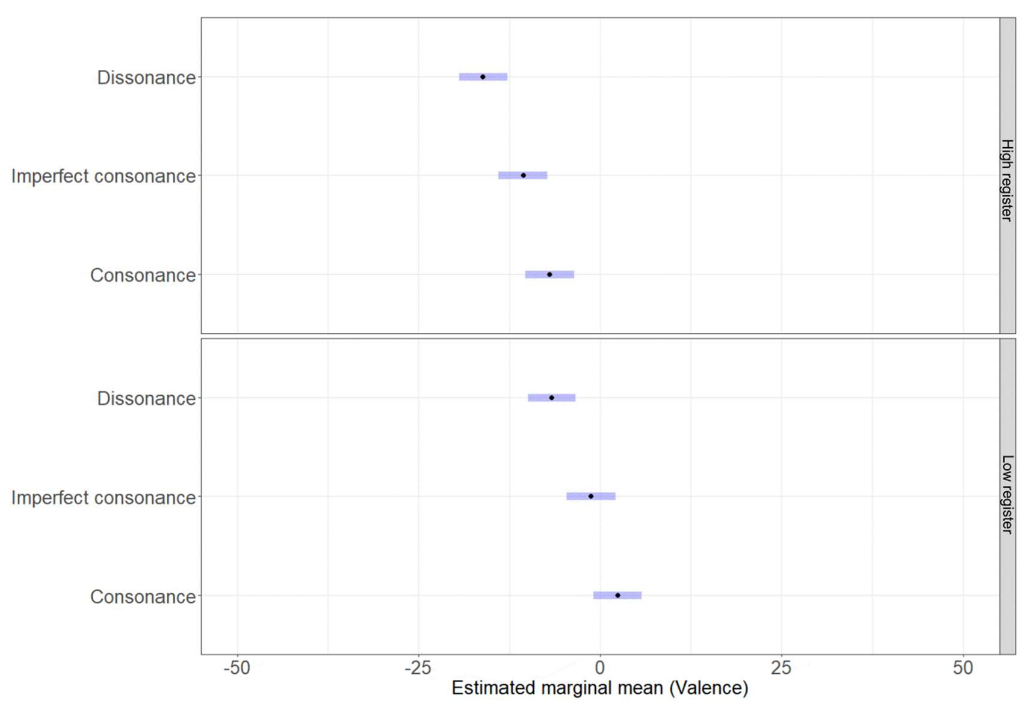
Table 6

Parameter Estimates of the Linear Mixed Model Analysis of Valence Rating of Harmonic Intervals

Predictor	*B*	*SE*	*t*	*p*
(Intercept)	-4.58	2.34	-1.93	.05
Roughness	-0.001	0.004	-1.63	.10
Imperfect consonance	-3.36	0.88	-3.80	< .001
Dissonance	-8.97	0.82	-10.88	< .001
Low register	10.53	0.96	10.90	< .001
Proprioceptive device	-4.73	3.06	-1.54	.13

Figure 12

Estimated Marginal Means and 95% Confidence Intervals for Musical Intervals Valence Ratings as a Function of Pitch Register



Tension – Musical intervals and noises

Table 7 shows the results of the mixed linear model analysis applied to tension ratings. Fixed factors were added sequentially, testing the significance of each *n* model with the *n -1* model. Significant effects resulted from the stimulus (musical interval and noises) and crossmodal procedure. Therefore, stimulus and crossmodal procedure were included in the final model, and the estimated parameters are shown in Table 8. Figure 13 shows the Q-Q plot for residuals. Estimated marginal means and 95% confidence intervals for tension ratings as a function of the crossmodal procedure are reported in Figure 14.

Table 7

Linear Mixed Model Results for the Tension Ratings

Model	*df*	*AIC*	χ^2	*p*
1. Intercept	3	22581.64		
2. 1 + Roughness	4	22580.18	3.45	.06
3. 2 + Years of musical study/practice	5	22582.06	0.12	.75
4. 3 + Crossmodal procedure	6	22573.27	10.79	.001
5. 4 + Stimulus	23	22253.20	354.06	< .001

Note. The fixed factors were sequentially included in the model. Participant was considered a random factor.

Table 8

Parameter Estimates in the Linear Mixed Model for Tension Ratings Including Crossmodal Procedure and Stimulus as Predictors

	B	*SE*	*t*	*p*
(Intercept)	46.71	2.84	16.40	< .001
Proprioceptive procedure	-10.68	3.08	-3.45	.001
m2	23.22	2.68	8.66	< .001
M2	13.46	2.68	5.02	< .001
m3	9.80	2.68	3.65	< .001
M3	2.98	2.68	1.11	.26
P4	2.99	2.68	1.11	.26
A4	14.90	2.68	5.56	< .001
P5	2.47	2.68	0.92	.35
m6	9.04	2.68	3.37	< .001
M6	-0.75	2.68	-0.28	.77
m7	9.84	2.68	3.67	< .001
M7	14.74	2.68	5.50	< .001
P8	2.46	2.68	0.91	.35
Pink	-10.56	3.28	-3.21	.001
Brown	-19.24	3.28	-5.86	< .001
White	-8.16	3.28	-2.48	.01
Blue	-0.79	3.28	-0.24	.81
Purple	5.09	3.28	1.55	.12

Figure 13

Q-Q Plot for the Linear Mixed Model Referred to Tension Rating

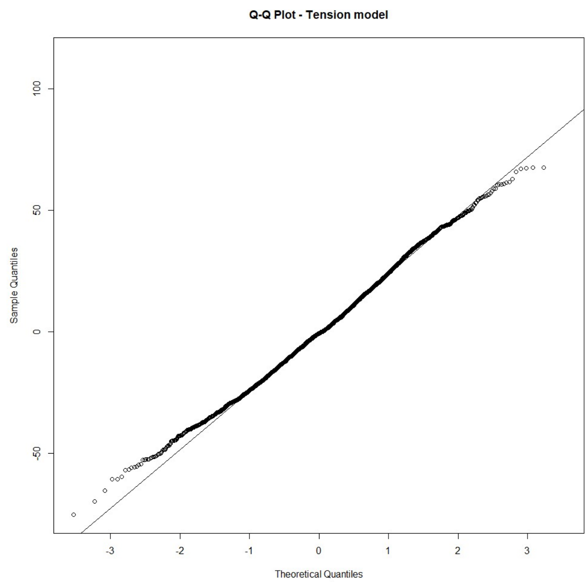
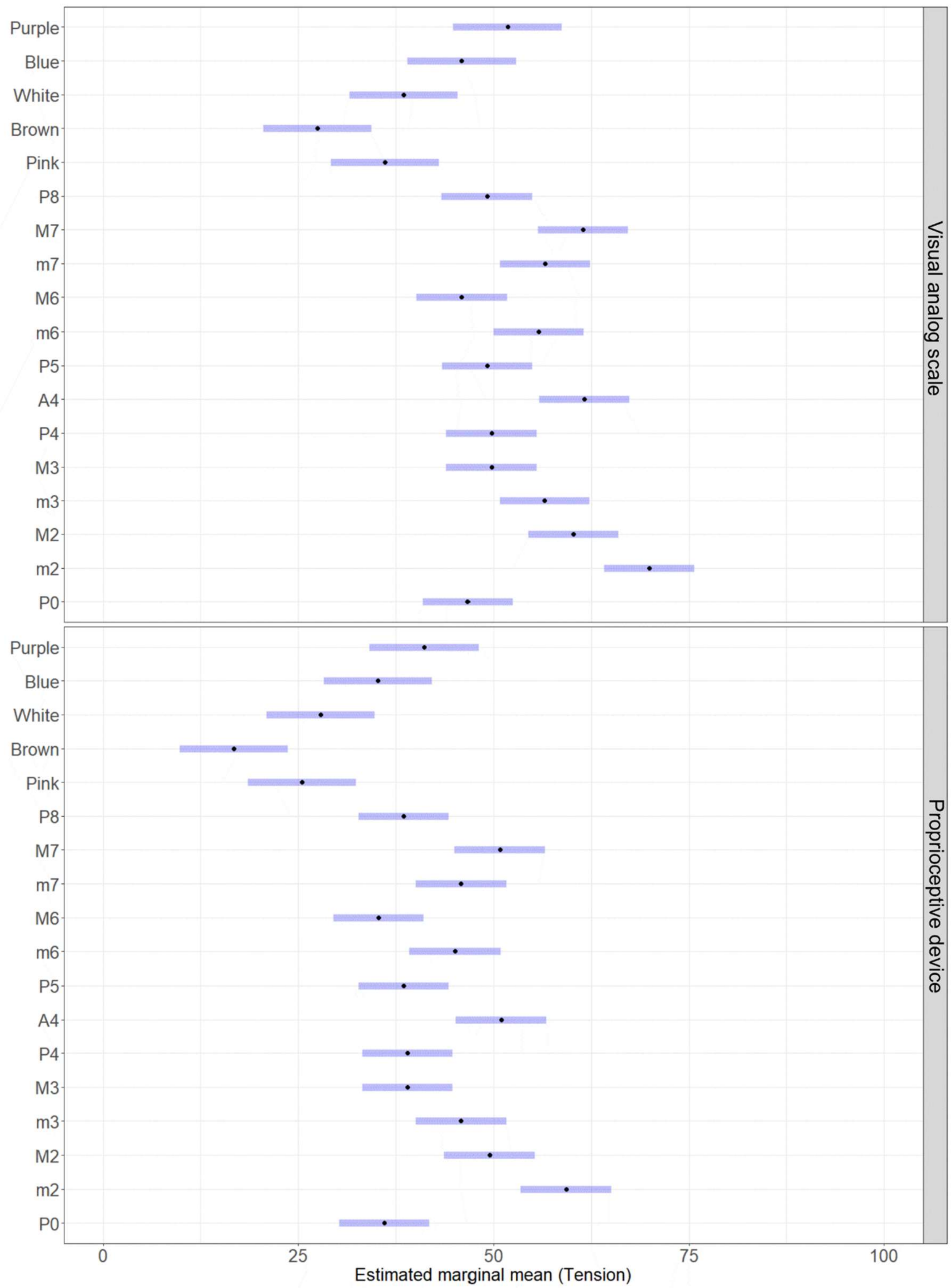


Figure 14

Estimated Marginal Means and 95% Confidence Intervals for Tension Rating as a Function of the Crossmodal Procedure



Tension – Musical intervals only

The linear mixed model for tension rating of musical intervals, including the level of consonance, register, roughness, and crossmodal procedure as predictors, showed that all the fixed effects and the covariate were significant, as shown in Table 9. Estimated marginal means and 95% confidence intervals for the three levels of consonance as a function of register and crossmodal procedure are shown in Figures 15 and 16, respectively. Tension received the highest ratings for dissonant intervals (*EMM*: 56.6, *SE*: 1.88), intermediate ratings for imperfect consonances (*EMM*: 46.3, *SE*: 1.93), and the lowest ratings for consonant intervals (*EMM*: 43.8, *SE*: 1.93). Tukey HSD tests showed that all the contrasts between the three levels were significant. Tension was evaluated higher for high-pitch register intervals (*EMM*: 54.9, *SE*: 1.91) in comparison to low-pitch register intervals (*EMM*: 42.8, *SE*: 1.91): $z = -8.67$, $p < .001$. Concerning the crossmodal procedure, the tension was rated higher when rated using the VAS (*EMM*: 53.7 *SE*: 2.51) than when rated with the proprioceptive device (*EMM*: 44.1, *SE*: 2.51): $z = -2.71$, $p = .006$.

Table 9

Parameter Estimates of the Linear Mixed Model Analysis of Tension Ratings of Musical Intervals

Predictor	*B*	*SE*	*t*	*p*
(Intercept)	52.62	2.65	19.81	< .001
Roughness	0.002	0.007	2.99	.002
Imperfect consonance	2.49	1.27	1.95	.05
Dissonance	12.84	1.19	10.78	< .001
Low register	-12.09	1.39	-8.65	< .001
Proprioceptive device	-9.63	3.55	-2.71	.01

Figure 15

Estimated Marginal Means and 95% Confidence Intervals of Tension Ratings as a Function of Consonance Level and Pitch Register

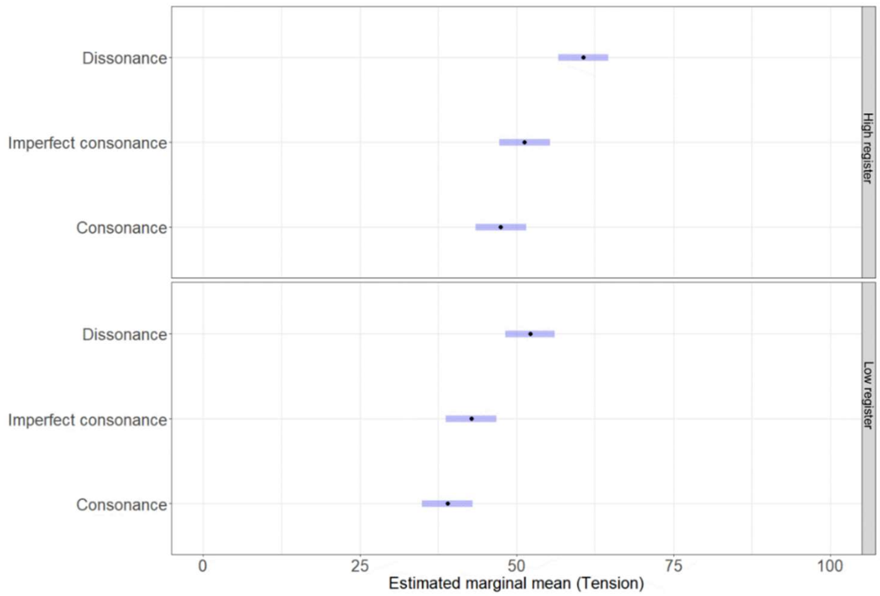
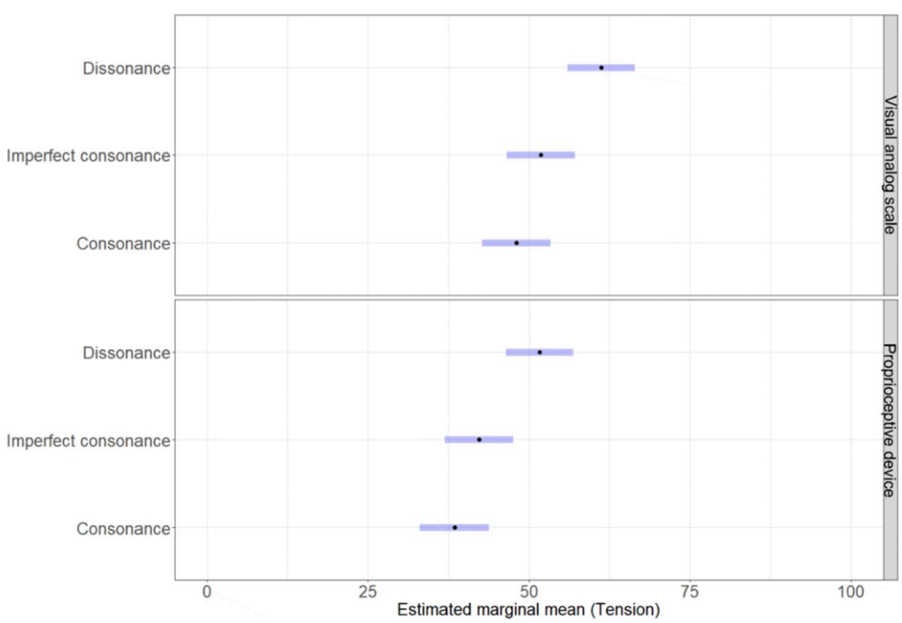


Figure 16

Estimated Marginal Means and 95% Confidence Intervals of Tension Ratings as a Function of Consonance Level and Crossmodal Procedure



Movement – Musical intervals and noises

Table 10 shows the results of the mixed linear model analysis of perceived movement ratings. We sequentially added each fixed factor, testing the significance of each *n* model with the *n -1* model. Stimulus (musical interval and noises) and roughness resulted in significant effects. Stimulus and roughness were, therefore, included in the final model. The estimated parameters are shown in Table 11, whereas Figure 17 shows the Q-Q plot for residuals. The estimated marginal means for perceived movement ratings are reported in Figure 18.

Table 10

Linear Mixed Model Results for Perceived Movement Ratings

Model	*df*	*AIC*	χ^2	*p*
1. Intercept	3	22470.76		
2. 1 + Roughness	4	22449.69	23.07	< .001
3. 2 + Years studying music	5	22450.77	.91	.33
4. 3 + Crossmodal procedure	6	22450.59	2.18	.13
5. 4 + Stimulus	22	22316.76	165.82	< .001

Note. The fixed factors were sequentially included in the model. Participant was considered a random factor.

Table 11

Parameter Estimates in the Linear Mixed Model for Movement Ratings Including Roughness and Stimulus as Predictors

	B	*SE*	*t*	*p*
(Intercept)	30.94	2.62	11.78	< .001
Roughness	-0.002	0.01	-3.94	< .001
m2	15.73	2.72	5.77	< .001
M2	7.08	2.79	2.53	.01
m3	12.09	2.77	4.34	< .001
M3	10.89	2.80	3.88	< .001
P4	9.30	2.80	3.31	< .001
A4	10.58	2.75	3.84	< .001
P5	7.61	2.70	2.80	.005
m6	9.15	2.71	3.37	< .001
M6	10.72	2.70	3.95	< .001
m7	10.15	2.69	3.76	< .001
M7	11.17	2.69	4.13	< .001
P8	4.91	2.69	1.82	.06
Pink	21.21	3.29	6.43	< .001
Brown	29.98	3.31	9.05	< .001
White	23.42	3.50	6.68	< .001
Blue	34.41	4.63	7.42	< .001
Purple	63.85	10.75	5.93	< .001

Figure 17

Q-Q Plot for the Linear Mixed Model Referred to Movement Rating

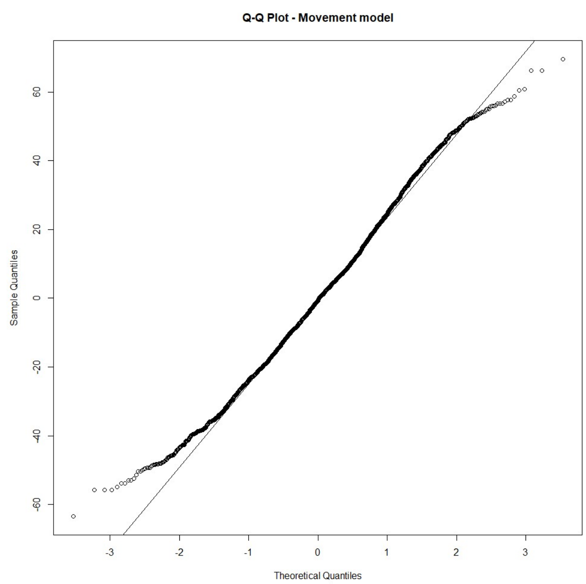
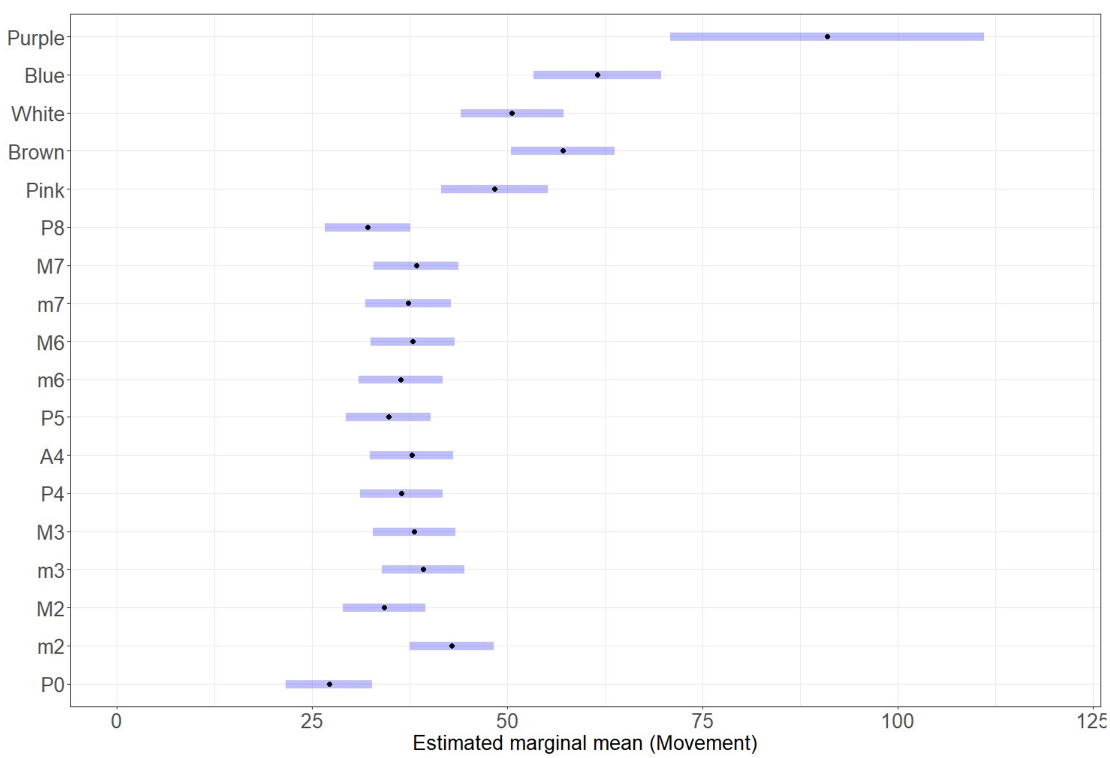


Figure 18

Estimated Marginal Means and 95% Confidence Intervals for Movement Rating



Movement – Musical intervals only

The linear mixed model for movement rating of musical intervals, including the level of consonance, register, roughness, and crossmodal procedure as predictors, showed that the fixed effects of consonance level and register were significant, as shown in Table 12. Estimated marginal means and 95% confidence intervals for the three levels of consonance as a function of pitch register are shown in Figure 19. The perception of movement was higher for dissonant intervals (*EMM*: 39.7, *SE*: 1.95) and imperfect consonances (*EMM*: 39, *SE*: 1.99) in comparison to consonant intervals (*EMM*: 34.8, *SE*: 1.99). Tukey HSD tests showed that the only contrast that was not significant was the difference between dissonant and imperfect consonant intervals. The perception of movement was significantly higher for high-register intervals (*EMM*: 41.3, *SE*: 1.98) than for low-register intervals (*EMM*: 34.3, *SE*: 1.98).

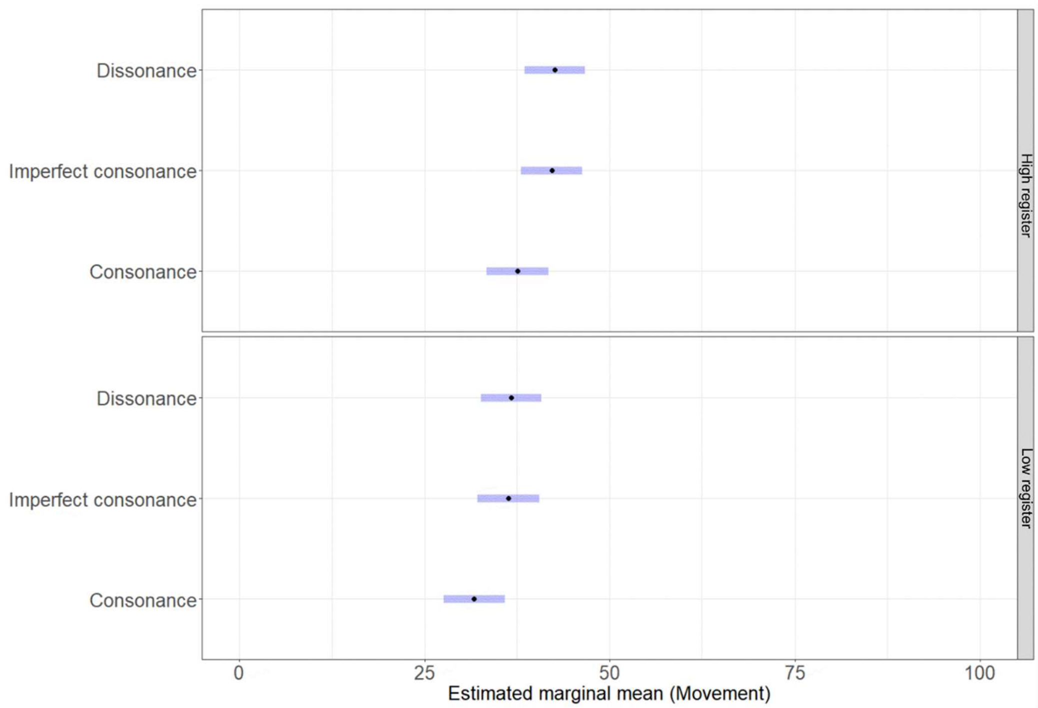
Table 12

Parameter Estimation of the Linear Mixed Model Analysis of Tension Rating in Musical Intervals

Predictor	*B*	*SE*	*t*	*p*
(Intercept)	40.41	2.75	14.66	< .001
Roughness	0.008	0.007	0.87	.38
Imperfect consonance	4.15	1.26	3.28	.001
Dissonance	4..91	1.18	4.16	< .001
Low register	-7.00	1.38	-5.05	< .001
Proprioceptive device	-5.33	3.70	-1.44	.15

Figure 19

Estimated Marginal Means and 95% Confidence Intervals for Perceived Movement Ratings



2.4 Discussion

This paper had five main goals: (a) to investigate the perception of tension, movement, and valence in musical intervals and specific standard noises; (b) to test the influence of roughness in the perception of tension, movement, and valence; (c) to assess the influence of pitch register on these three attributes referred to musical intervals only; (d) to compare tension, movement, and valence between 'voiced' musical intervals and 'unvoiced' standard noises that differed in their spectral emphasis of low or high frequencies; (e) to compare two crossmodal methods for the assessment of the perception of tension and movement: one that mapped tension and movement along with a VAS and the other that relied on a proprioceptive device in which tension and movement were mapped with muscular force and pulling angle.

As expected, the three attributes of valence, tension, and movement were not independent but showed a discrete degree of independence. The strongest (inverse) relationship emerged between valence and tension. Tension was mainly perceived with sounds that were judged as unpleasant and with a negative valence. This aspect is interesting in the perspective of music and emotion theory. Some authors have associated tension with affective arousal (Krumhansl, 1997; Trolio, 1976). Rozin et al. (2004), for example, measured moment-to-moment 'affective intensity,' and Huron's model of expectation (2006) includes an arousal-related tension component. The association of tension with arousal tends to be further promoted by the common use of a 2D arousal-valence space for collecting data on emotional responses to music.

Other authors proposed a need to differentiate between tension and energy arousal. Specifically, Thayer (1989) reconceptualized activation as varying along two dimensions: energetic arousal (awake-tired) and tense arousal (tense-calm), and this distinction was further employed by Ilie and Thompson (2006) and Schimmack and Rainer (2002). Eerola and Vuoskoski (2011) tested a 3D model for emotion in music that included valence, tension, and energy as key dimensions.

In this study, tension and valence were not completely orthogonal factors since tension was mainly associated with dissonant intervals. The highest linear coefficients in the tension model were found for the minor second, the augmented fourth, and major seventh, which are the most dissonant intervals according to music theory.

A similarly inversely related pattern between valence and tension, mediated by the spectrum-related energy content, was found for standard noises. The noises that emphasized high-frequencies (i.e., purple and blue) were judged extremely negatively for valence compared to the brown and pink noises whose spectral density emphasizes low frequencies. When considering tension, pink and brown noises were judged less tense than blue and purple noises. Interestingly, while purple and blue noises were judged as the most unpleasant among all the stimuli, the pattern was not symmetrical when observing perceived tension; in fact, tension ratings of these two noises were lower than those of the most dissonant intervals (minor second, augmented fourth and major seventh).

Perceived tension was strongly related to pitch register both in musical intervals and noises. Perceived tension was lower in the low-pitch intervals and noises that emphasized low-frequencies (i.e., pink and brown noises). This outcome is in line with previous studies where low-pitched music was rated as more pleasant and less tense (Ilie & Thompson, 2006), and with the model proposed by Farbood (2012), who found that sequences of descending chords were mostly associated with a decrease in tension and sequences of ascending chords were associated with an increase in tension. The use of more structured stimuli in these other studies implied that pitch register was not isolated from the melodic contour. Low-pitch stimuli were also often the results of a descending melodic line and, since pitch is strongly mapped on a vertical space (Bonetti & Costa, 2018), the decrease in tension could have been the result of a perceived descending melodic line. Since the stimuli were stationary and composed of steady musical intervals and noises and the pitch register effect was very pronounced in the present study, we can conclude that this factor is one of the best predictors of perceived tension. McAdams, Douglas, and Vempala (2017) investigated the perception of affective qualities of musical instrument sounds across the pitch register and found the same result, and also that higher tension was carried by brighter sound.

This association between high pitch and an increase in perceived tension could be explained in an ecological-evolutionary framework. From this perspective, musical tension is affected by those auditory features associated with tension in ‘natural’ extramusical contexts. An increase in pitch-height in vocal emissions is a signal of distress, fear, anger, and isolation in many species. A significant increase in pitch characterizes most alarm calls used in social animals to alert conspecifics about a predator's presence compared to normal vocalizations

(Fallow et al., 2011). High-pitch vocalizations are reliably perceived as an indicator of distress in infant cries (Schuetze & Zeskind, 2001; Soltis, 2004; Zeskind & Marshall, 1988), and an increase in pitch in the voice is frequently associated with the experience of distress and tense emotions as fear, anger (Sobin & Alpert, 1999).

Starting with Helmholtz (1887), the beatings of adjacent partials have been one of the main explanatory factors for dissonance perception. Plomp & Levelt (1965) have further developed this theory introducing the notions of critical bandwidth and sensory consonance (Terhardt, 1978), thus distinguishing the consonance due to basic physical and physiological factors from the consonance influenced by more high-order factors. The term 'roughness' was then preferred over the expression 'sensory consonance' because it could also be applied to amplitude-modulated tones (Zwicker & Faswtl, 1990). In the present study, the role of roughness, computed according to the model of Sethares (2005), was tested in the perception of valence, tension, and movement, introducing a comparison between musical intervals and standard noises. By definition, standard noise comprises all frequencies in the range of acoustical perception (typically 20 – 20,000 Hz). The energetic content of each frequency is regulated by a mathematical function, and in our case, we choose five noises that differed in their emphasis on low- and high-pitch frequencies. Specifically, there was an emphasis on low-pitch frequencies in brown and pink noises and high-pitch frequencies in blue and purple noise. In white noise, the energy amplitude was flat all over the frequency range. The roughness level in noises is strongly influenced by the content of high-pitch frequencies, reaching a peak in purple noise, whose roughness level was 12.20 times greater than that of the minor second built on C3 and 37.98 times greater than that of the minor second built on G4.

Nevertheless, the results showed that the rated tension for purple noise was lower than that attributed to the dissonant musical intervals of minor second, augmented fourth, and major seventh. The linear mixed model results that included both musical intervals and noises showed that roughness was not a significant predictor of perceived tension, while it was a significant predictor in the case of valence. Therefore, it seems that the relation between roughness and tension is not straightforward, as in the case of pleasantness/unpleasantness. For example, low-register musical intervals had a significantly higher level of roughness than high-register intervals, but the perception of tension was opposite, with high-register intervals perceived as tenser than low-register intervals.

Considering musical intervals, tension was proportional to the level of dissonance of the interval. The intervals that were perceived as tenser were the seconds, seventh, and augmented fourth. The rank order of tension perception in intervals strictly mirrored the rank order of intervals by consonance and dissonance in the classic study by Malmberg (1918). A similar effect applied to single chords was also found by Lahdelma and Eerola (2016a, 2016b). They found that perceived tension was very high for the Neapolitan pentachord, followed by the dominant seventh sharp eleventh chord. The lowest tension level was found for the major triad, which was the most consonant chord in their study. Tension for minor chords was higher than tension for major chords. Tension was also affected by the chord's position, increasing linearly from the root position to the first and second inversion. In another study using chords, the same authors also found a high correlation between tension and energy (.50) (Lahdelma & Eerola, 2016b); augmented and diminished chords elicited the highest tension ratings, followed by sevenths and minor and major chords.

The comparison between the two crossmodal procedures used to assess tension and movement in Experiments 2 and 3 showed a significant difference only in the evaluation of tension. Specifically, the use of the proprioceptive device led systematically to tension evaluations that were lower than those obtained with the VAS. The distribution related to the proprioceptive device was positively skewed, while the VAS rating distribution was negatively skewed. In Experiment 1, the proprioceptive device was tested for a linear psychophysical relation between applied force and perceived force. Thus, the distribution obtained in Experiments 2 and 3 could not be attributed to an intrinsic non-linearity in force perception along its angular displacement. The effect could be due to the proprioceptive task involving a matching between perceived tension and muscular tension. In this case, the response implied more effort and more direct feedback on the increase in the tension level, whereas, in the VAS, the mapping of tension on the horizontal line could have been less steep.

The crossmodal procedure did not alter the pattern of perceived tension as a function of the stimuli presented in Experiments 2 and 3. The pattern remained substantially the same and was simply shifted between the two procedures. The results cannot favor one procedure over the other but are important in showing that tension ratings are highly susceptible to the methodology used for collecting the data, and a standardization procedure would be preferable to relying on absolute values.

Roughness was a strong predictor in the perception of movement, but only for standard noises. Blue and purple noises, which shared a high spectral content of high frequencies, were perceived as inducing a higher sense of movement than white, pink, and brown noises. The linear mixed model analysis showed that the strongest predictors of movement ratings were pitch register (movement perception was higher for high-register intervals) and the level of dissonance of the interval.

In conclusion, although related by a certain degree of commonality, the attributes of valence, tension, and movement applied to musical intervals and noises appear to have distinctive properties that cannot be reduced to a single core explained by roughness level. For example, low-register intervals had a higher level of roughness, but they were perceived as more pleasant, less tense, and inducing a lower sense of movement than high-register intervals. Roughness was a significant predictor of valence and movement, but not for tension when including standard noises in the analysis. The effect of pitch register reflected shared between the attributes of valence, tension, and movement; its influence was consistent between the three domains. The intervals and noises perceived as more unpleasant were also perceived as tenser and inducing a high sense of movement. This outcome has interesting implications, especially for the status of tension in the theory of music-evoked emotions. Specifically, these results question the complete orthogonality of tension, conceptualized as an arousal component distinct from valence. Although it is certainly possible to conceive states of high tension in conjunction with positive valence due to high-order musical elements, for example, a crescendo and accelerando of a major mode melody, in the basic vocabulary of musical intervals and standard noises, tension is solely associated with the experience of negative valence. In contrast, positive valence is associated with the perception of release and steadiness.

This paper has also shown how the attributes of valence, tension, and movement applied to musical intervals can be applied to unvoiced acoustical stimuli as noises. Standard noises are interesting research tools because they can physically be considered supernormal stimuli of dissonant intervals. For example, their ratings of perceived movement exceeded those of dissonant musical intervals, although some distinctive properties emerged in the case of tension and valence. For example, Brown noise was evaluated as the less tense stimulus, even less than the perfect consonant intervals (i.e., the unison and the octave). Similarly, for valence, brown noise had the highest ratings for pleasantness. The combined study of musical intervals and noises could significantly contribute to shed light on the processes that underlie the attribution of psychological qualities to sounds.

3 Neural processing of tension-resolution patterns based on consonance-dissonance

3.1 Introduction

Music listening is a dynamic experience that involves the processing of discrete elements combined in more complex structures unfolding through time. This flux inherent to music is also a key element in the perception of musical tension, which arises from the combined interaction of various musical elements. Local and simple tension–resolution patterns are organized in a hierarchical structure forming a global and complex tension–resolution pattern. Numerous tension arches are usually intertwined into large-scale tension arches in Western music. For example, single notes can be combined into a melody to form a musical phrase. Such a phrase can constitute a distinct tension arch built on syntactic rules that create expectations in listeners.

Nonetheless, more phrases can be linked together to create an overlying larger tension arch. These hierarchical structures of musical dependencies can grow to the ultimate level, represented by the entire composition. These concepts have been described in the generative theory of tonal music (GTTM; Lerdahl and Jackendoff, 1983) and the tonal tension model (TTM; Lerdahl and Krumhansl, 2007).

Several researchers tried to highlight the neural mechanisms underpinning the perception of musical tension using different techniques. Some of these techniques allow us to observe neural changes associated with music listening of relatively long music excerpts (e.g., fMRI, qEEG), while others allow analyzing quick changes that occur at specific time points within relatively short music excerpts (e.g., ERPs, MEG).

These techniques present deep methodological differences and usually answer different research questions. A list of studies employing the ERP technique to study aspects linked to musical tension is provided below. Since the present study procedure relies upon consonant-dissonant harmonic intervals to create tension-resolution patterns, studies focusing on sensory consonance-dissonance will also be summarized. Although some of these studies did not directly use the term ‘tension’ within their papers, they employed common tension eliciting procedures (i.e., harmonic violations).

Regnault et al. (2001) recorded the changes in the brain electrical activity associated with the presentation of chord sequences to analyze the time-course of sensory (bottom-up) and cognitive (top-down) processes that govern musical harmonic expectancy. Stimuli consisted of eight-chord sequences where the expectation on the final chord (target) was manipulated both at the sensory level (i.e., the last chord was sensory consonant or dissonant) and at the cognitive level (i.e., the harmonic function of the target was varied by manipulating the harmonic context built up by the first six chords of the sequence). Changes in the harmonic context modulated the amplitude of the P3 component, reflecting top-down influences on the perceptual stages of processing. In contrast, changes in the acoustic structure of the target chord (sensory consonance) produced a larger late positive component (LPC; Besson & Faïta, 1995) between 300 and 800 ms after target onset dissonant targets compared to consonant ones. These two effects (sensory consonance and harmonic context) were independent, suggesting that two separate processors contribute to the building up of musical expectancy.

Koelsch et al. (2007) investigated music-syntactic processing with chord sequences that ended on either regular or irregular ways based on harmonic regularities. Sequences were composed in order to disentangle the cognitive processing of syntactically regular and irregular chords and the sensory processing of acoustic factors like horizontal sensory dissonance or roughness. Irregular chords elicited an early right anterior negativity (ERAN; peaking around 200 ms from the onset of the last chord) in the ERPs under both task-relevant and task-irrelevant conditions. Notably, this effect seemed independent from the subjective detection of irregularities based on acoustical features such as pitch commonality and roughness. In this study, other components were elicited by syntactic irregularities. A larger N5 component (maximal around 500–550 ms) was recorded after irregular chords, possibly reflecting processes of harmonic integration (Koelsch et al., 2000). The authors suggested that while regular final chords can easily be integrated into the established musical context, irregular chords require a larger amount of harmonic integration reflecting in a larger N5. However, the exact relation between N5 and processing of musical meaning remains to be specified. Also, an LPC was elicited by irregular chords when these were task-irrelevant. This component may reflect structural integration processes and possibly processes of structural repair (Koelsch & Siebel, 2005).

Steinbeis and Koelsch (2008) combined syntactic and semantic linguistic violations with harmonic expectation violations in tension-resolution musical patterns created using five-chord sequences. An early right anterior negativity (ERAN) and the N500 were systematically modulated; the ERAN by the concurrent presentation of musical stimuli with a syntactic language violation; the N500 by the concurrent presentation of musical stimuli with a semantic language violation. This study showed that tension-resolution patterns represent a route to meaning in music and that the N500 can be therefore interpreted as reflecting the processing of semantic aspects of tension-resolution patterns.

The studies mentioned above systematically manipulated the syntactic aspects of music in order to create tension evoking patterns related to the violation of expectation. Other studies focused on the neural correlates of more basic acoustic properties related to harmony (e.g., consonance/dissonance).

Schön et al. (2005) conducted a series of experiments to determine whether consonant and dissonant intervals out of any musical context elicit similar or different neurophysiological responses and whether these effects are similar or different for musicians and nonmusicians.

They recorded the ERPs elicited by the different intervals classified according to music theory into three categories: perfect consonances, imperfect consonances, and dissonances. Participants were also required to rate, on a six-point scale, whether the intervals evoked pleasant or unpleasant feelings. Two notes were either played together (harmonic intervals) or successively (melodic intervals). In nonmusicians, a larger N2 component for harmonic consonance than for imperfect consonances and dissonances. A larger N420 component was elicited by imperfect consonances rather than by perfect consonances and dissonances. When the results were analyzed as a function of pleasantness responses, the N1-P2 complex was more positive for the intervals judged as very pleasant than for those judged as unpleasant. Both for harmonic and melodic intervals, a significant main effect of response type was found starting at a 420 ms latency and lasting for ~200 ms in harmonic intervals and for ~100 ms in the case of melodic intervals.

Itoh et al. (2003) investigated the neural processing underlying perception of noncontextual musical consonance using ERP. Participants listened to a random sequence of dyads formed on five different harmonic intervals (m2, M3, A4, P5, or M6). Amplitudes of P2 and N2 components of auditory ERPs were significantly modulated by the distance between the two notes of the dyad, in other words, by the type of interval. The largest negative amplitude was observed for the minor second (m2) while the least negative (or most positive) amplitude for the perfect fifth (P5) for both P2 and N2. The results indicate that the P2 and N2 components are affected by the harmonic relationship between the two notes forming the dyads.

The same authors (Itoh et al., 2011) replicated and extended these results using 14 intervals (from the unison to the minor ninth). Previous results were confirmed again, highlighting the N2 and P2 components as relevant for the processing of harmonic relationships between two notes. Further, this effect was evident only in musicians, indicating that there is plasticity in the neural processing of noncontextual consonance. These results support the hypothesis that the central auditory processing of properties of intervals other than those related to roughness contributes to the noncontextual consonance perception.

Proverbio et al. (2016) investigated how musical expertise influences the auditory processing of harmonicity. They presented musicians and nonmusicians with 200 chords consisting of pure tones, manipulating the degree of consonance-dissonance. An early auditory N1 was observed that was modulated by chord dissonance in both groups. An anterior negative component (N2) was enhanced only in musicians in response to chords featuring quartertones, suggesting a greater pitch sensitivity for simultaneous pure tones in the skilled brain. The P300 was affected by the frequency range only in musicians, who also showed a greater sensitivity to sound complexity. A strong left hemispheric specialization for processing quarter tones in the left temporal cortex of musicians was observed at N2 level (250–350 ms), which was observed on the right side in controls.

Crespo-Bojorque et al. (2018) conducted an ERP study using an oddball paradigm to highlight the passive processing of consonant and dissonant music intervals. Participants were presented with sequences of consonant intervals interrupted by a dissonant interval or sequences of dissonant intervals interrupted by a consonant interval. They found that brain responses elicited for the transition from consonance to dissonance differ from brain responses to the transition from dissonance to consonance. Changes in a sequence of consonant intervals elicited a mismatch negativity (MMN) between 150 ms and 250 ms from the stimulus onset both in musicians and nonmusicians, independently of music expertise. In contrast, changes in a sequence of dissonant intervals elicited a late MMN only in participants with prolonged musical training. These different neural responses might form the basis for the processing advantages observed for consonance over dissonance and provide information about how formal musical training modulates them. Interestingly, musician participants also exhibited a significant negative response from 502 to 582 ms (N500). As already mentioned, this component can be elicited by violations of harmonic expectations (Koelsch et al., 2000; Koelsch & Siebel, 2005).

Previous studies investigated music processing focusing on the neural correlates of either noncontextual consonance (Crespo-Bojorque et al., 2018; Itoh et al., 2003, 2011) or violations of expectations built on the tonal context given by chords sequences (Koelsch et al., 2007; Regnault et al., 2001; Steinbeis & Koelsch, 2008).

The present experiment aimed to bridge this gap by eliciting tension and resolution using a minimal harmonic context with the least syntactic complexity. Short tension-resolution patterns based on sensory consonance-dissonance were presented in a tension-detection task while recording brain electrical activity to understand how consonance-dissonance can convey meaningful information of perceived tension-resolution. These patterns were formed by sequences of two adjacent harmonic intervals, the first consonant and the second dissonant, to create the perception of tension ('tension induction' condition), and vice-versa to create a resolution perception ('resolution' condition). Intervals falling at the extreme points of the consonance-dissonance continuum were chosen based both on music theory and the results of the previous experiment. Two other experimental conditions were created, containing two intervals that were both consonant or dissonant, to disentangle the processing of musical tension from the sensory processing of consonance-dissonance.

Because of its excellent temporal resolution, the ERP method allows us to precisely determine at which point in time the neural signals elicited in two experimental conditions start to diverge, and by inference, to determine when the two underlying processes start to differ.

Based on the mentioned results, early components (N2, P2) related to the processing of harmonic properties are expected to be influenced by the harmonic dependency of the second interval from the first interval, as well as late components (N500, LPC) related to the contextual/syntactic processing of musical stimuli.

3.2 Methods

Participants

Sixteen volunteers (11 females, mean age 71.31 ± 3.56) participated in the experiment. All participants took part in a larger study investigating the effects of musical training on brain plasticity in the elderly and therefore received one year of music training. In particular, seven participants received piano lessons, and nine received music culture lessons. None of them had received music training before taking part in the study. All participants were right-handed, reported normal hearing, signed written informed consent, and received monetary compensation for their participation in the study. A posthoc power analysis was performed using G∗power (v. 3.1.9.4), indicating a sample of 16 and eight levels of measurement corresponding to the within factors of a repeated-measures ANOVA. The results showed that a 90% power was obtained for detecting an effect size of $f = 0.25$, with $\alpha = 0.05$.

Materials and experimental procedure

Stimuli consisted of three consonant and three dissonant harmonic intervals (two-note chords). Consonant intervals were the perfect unison (P0), the perfect fourth (P4), and the perfect fifth (P5). Dissonant intervals were the augmented fourth (A4), the minor second (m2), and the inverted major seventh based on the lower octave (iM7), which corresponds to an interval composed by the root note and a note one semitone lower than the root note (Table 13). Intervals were created using the software Csound with a flute timbre using the 5-limit tuning just ratios to calculate the exact frequency of the upper tones. Loudness was equalized to 12 sones with the Matlab Genesis loudness Toolbox (Genesis, 2009), applying the ANSI S34 2007 procedure (American National Standards Institute, 2007). These intervals were then combined, creating two-interval sequences.

Table 13

Consonant and Dissonant Harmonic Intervals Used in the Present Study

	Stimuli		Key		
	Interval	Ratio	F	C	G
Consonant	P0	1:1	F_3-F_3	C_4-C_4	G_4-G_4
	P4	4:3	F_3-$B\flat_3$	C_4-F_4	G_4-C_5
	P5	3:2	F_3-C_4	C_4-G_4	G_4-D_5
Dissonant	A4	45:32	F_3-B_3	C_4-$F\sharp_4$	G_4-$C\sharp_5$
	m2	16:15	F_3-$G\flat_3$	C_4-$D\flat_4$	G_4-$A\flat_4$
	iM7	15:16	F_3-E_3	C_4-B_3	G_4-$F\sharp_4$

Note. A lower (the key) and an upper note with a correspondent frequency ratio compose the interval. Each interval used in the present experiment was implemented in three different keys: F, C, and G. The corresponding upper notes are specified for each interval.

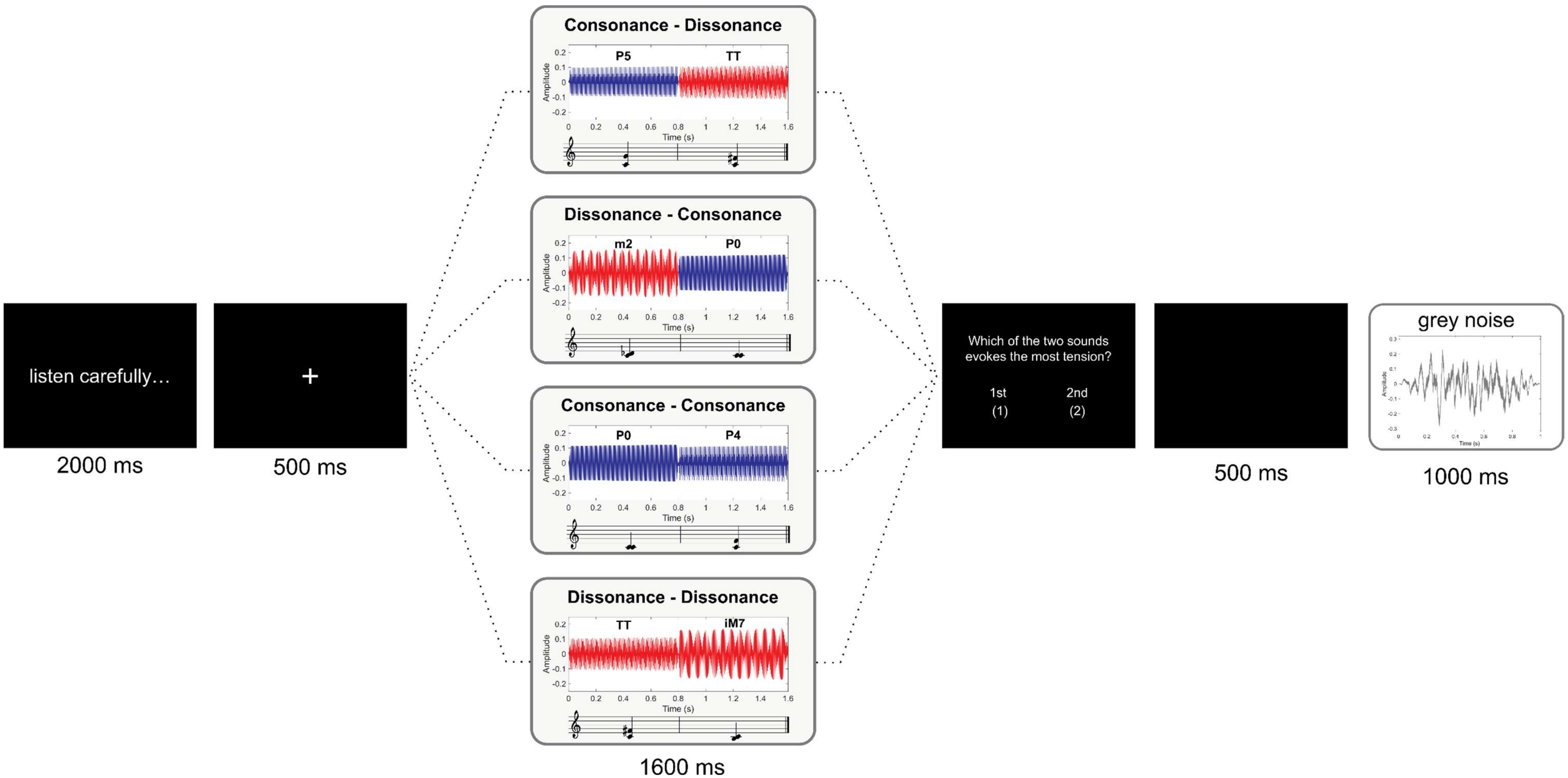
Four conditions based on the order of consonant and dissonant intervals were created: consonance – dissonance (CD); dissonance – consonance (DC); consonance – consonance (CC); dissonance – dissonant (DD). The first two conditions contained intervals differing in their level of consonance/dissonance, while the other two contained intervals which were similarly consonant/dissonant. The consonance/dissonance varying sequences (CD and DC) represented tension-resolution patterns where the instability created by dissonance was resolved in more stable and consonant intervals and vice-versa (tension/resolution conditions), while the sequences with a constant level of consonance/dissonance were used as control conditions. CC and DD sequences were created using all the permutations without repetition of consonant or dissonant intervals. CD and DC sequences were created using contiguous intervals where the upper tone varied melodically by one semitone (ascending or descending) between the first and the second interval so that the tension between the two intervals was emphasized by the melodic attraction of the two upper tones. In order to generalize results to a wider pitch range, three different keys were used to create intervals (F3, C4, G4), so that sounds ranged from 163.701 to 587.993 Hz (E_3 to D_5). Each harmonic interval was 800 ms long with an onset/offset 10 ms ramp. Thus each two-interval sequence was 1600 ms long.

Participants were individually tested while seating comfortably in front of a computer screen in a soundproof room. For each two-interval sequence, they were requested to indicate which of the two intervals evoked more tension. The stimulus sequence is shown in Figure 20. At the beginning of each trial, a fixation cross appeared in the middle of the screen for 500 ms. Then, a sound sequence consisting of two harmonic intervals (800 ms each) was presented through two loudspeakers positioned at the sides of the screen. There was no silence between the onset of the first interval and the onset of the second one, nor was there any overlap. Immediately after, participants were prompted with the following question: “Which of the two sounds evokes the most tension?”. They were encouraged to respond naturally and not to spend much time on the decision.

They were also asked to respond only after the prompt appeared on the screen (at the end of the second interval). Using the right hand, they had to press key ‘1’ (index) if the tensest interval was the first, or key ‘2’ (middle finger) if the tensest interval was the second. A grey noise was presented for 1000 ms after the response acquisition to prevent the retaining of harmonic intervals in the auditory sensory memory, and a delay of 500 ms preceded the next stimulus. All participants were presented with all the four conditions, and each condition was presented 72 times, for a total of 288 trials. The experimental session was divided into four blocks of 72 trials (7-8 minutes/block). Within each block, the four conditions were equally represented (25%), and the order of stimuli was randomized.

Figure 20

Stimulus Sequence With Four Types of Two-Intervals Combinations



Note. Examples of sound waves and music notations are provided in C key for each condition. The Consonance-Dissonance condition corresponds to the 'Tension induction' pattern, whereas the Dissonance-Consonance condition corresponds to the 'Resolution' pattern.

There was a break of 3 min between blocks. The total duration of the experiment was around 45 min. The sound sequences were presented using the software ePrime 3.0. All experimental procedures were approved by the ethical committee of the Geneva School of Health Sciences, University of Applied Sciences and Arts Western Switzerland (HES-SO). Participants were allowed to adjust the output of the amplifier to a comfortable level before starting the experimental session. A training session consisting of eight random trials was included to ensure the understanding of the experimental procedure.

EEG acquisition and raw data processing

EEG was continuously recorded from 64 electrode sites (BioSemi Active-Two, V.O.F., Amsterdam, the Netherlands), equally distributed over the scalp. Data were digitized at a sampling rate of 2,048 Hz in a bandwidth filter of 0–268 Hz. Electrodes' impedances were kept below 20 Ω. Prior to analysis, latency was corrected by shifting the EEG data by 8 ms as this was the latency between the trigger delivery and the actual sound output estimated during a prior pilot test. Data were offline recomputed against the average reference, band-pass filtered (1–30 Hz) with a 2nd- order Butterworth filter (−12 dB/octave roll-off; Brunet et al., 2011) and downsampled to 256 Hz to speed up the next processing steps. A DC shift correction was applied. In addition to an automated threshold rejection criterion of 100 μV, epochs were visually inspected for oculomotor and other artifacts using preprocessing methods (Brunet et al., 2011; Luck, 2014). On average, 60 ± 11.2 (83.3 %) epochs per condition per participant were retained. Channels exhibiting substantial noise were interpolated using a 3D spherical spline interpolation (3.8% ± 3.6 of all channels) that accounts for the real geometry of the head (Brunet et al., 2011; Perrin et al., 1989). Averaged evoked potentials were calculated from -200 up to 800 ms after the stimulus onset (the second harmonic interval). For waveform analyses, ERPs were baseline corrected ('pre-stimulus' baseline from −200ms to stimulus onset) to allow comparison to the literature. This data processing was performed with the Cartool software, developed by Denis Brunet (brainmapping.unige.ch/cartool).

ERP analysis

For statistical evaluation, ERPs were analyzed by repeated-measures analyses of variance (ANOVAs) as univariate tests of hypotheses for within-subject effects. Mean ERP values were computed for four regions of interest (ROIs): left anterior (F7, F5, F3, FT7, FC5, and FC3), right anterior (F8, F6, F4, FT8, FC6, and FC4), left posterior (TP7, CP5, CP3, P7, P5, and P3), and right posterior (TP8, CP6, CP4, P8, P6, and P4). The factors entering the ANOVAs were the following: type of sequence (CD × DC), hemisphere (right × left), and antpost (anterior × posterior). The time windows used for the analyses were 60-100 ms, 200-300 ms, 300-400 ms, and 500-600 ms.

3.3 Results

Behavioral results

The mean response accuracy computed considering only the combinations composed of different categories of intervals (i.e., CD and DC) was $M = 0.73$.

The mean response accuracy in the two tension-varying conditions was compared to verify the presence of confounding effects across time (four blocks of stimuli), across pitch registers (three levels: F3, C4, G4), melodic contour (two levels: ascending, descending), the condition type (two levels: consonance-dissonance, dissonance-consonance), and the combination of intervals (eight levels: P0 – m2, P0 – iM7, m2 – P0, iM7 – P0, P4 – A4, P5 – A4, A4 – P4, A4 – P5). Only the combination of intervals significantly affected the mean accuracy; the Greenhouse-Geisser estimate of the departure of sphericity was $\varepsilon = .58$. The mean response accuracy was significantly affected by the sequence of stimuli presented, $F(4.05, 60.67) = 14.40$, $p < .001$, $\eta_p^2 = .49$. Mean and standard errors of accuracy for each combination are reported in Table 14.

Medians of individual response times were compared between the ‘tension induction’ condition and the ‘resolution’ condition only for correct responses using a paired-samples t test. Response time did not differ significantly between the ‘tension induction’ condition ($M = 514.63$, $SE = 49.84$) and the ‘resolution’ condition (*M* = 620.59, *SE* = 66.03), *t* (15) = -1.90, *p* = .077, BCa 95% CI [-221.39, -.98], *d* = .44 .

Table 14

Means and Standard Errors of Response Accuracy for Each Combination of Intervals

Condition	Stimuli	*M* (%)	*SE* (%)
Tension induction (CD)	P0 – m2	93.40	1.47
	P0 – iM7	87.15	1.98
	P4 – A4	59.03	2.90
	P5 – A4	56.25	2.93
Resolution (DC)	m2 – P0	84.03	2.16
	iM7 – P0	77.78	2.45
	A4 – P4	64.24	2.83
	A4 – P5	64.58	2.82

ERP results

The ‘tension induction’ condition elicited a significant main effect of condition (F(1, 15) = 7.24, p = .017) within the time window 300-400, indicating an increased positivity over the left frontal sites and a decreased positivity over the right frontal sites in the CD condition compared to the DD condition. Comparisons between signals within the other analyzed temporal windows did not show significant results.

The ‘resolution’ condition elicited a P1 component with a peak amplitude around 80 ms higher in the CC condition than the DC condition, mainly visible in the frontal sites. However, the ANOVA did not highlight a significant effect of the condition within this time window. A P2 component was visible with a peak around 250 ms, mainly in the frontal sites. A significant 2-way interaction effect with the factors condition and antpost (*F*(1, 15) = 4.63, *p* = .048) emerged within the time window

200 - 300 ms, indicating a decreased positivity over anterior sites than over posterior ones in the DC condition compared to the CC condition. A P300 component was evoked around 300 ms. A significant 2-way interaction effect with the factors condition and antpost ($F(1, 15) = 5{,}70$, $p = .031$) within the time window 300 - 400 ms, indicating a decreased positivity over anterior sites than over posterior ones in the DC condition compared to the CC condition. The DC condition also elicited a late positive component (LPC) frontal component, starting around 500 ms and lasting for about 100 ms. A main effect of condition ($F(1, 15) = 6.75$, $p = .02$) within the time window 500-600 was significant. A 2-way interaction ($F(1, 15) = 6.62$, $p = .021$) indicating an increased negativity over the anterior sites than over posterior ones in the CC condition compared to the DD condition.

Figures 21 and 22 show the averaged ERP for selected channels, and topographical activations emerged from the comparison of DD and CD condition ('tension induction') and CC and DC conditions ('resolution').

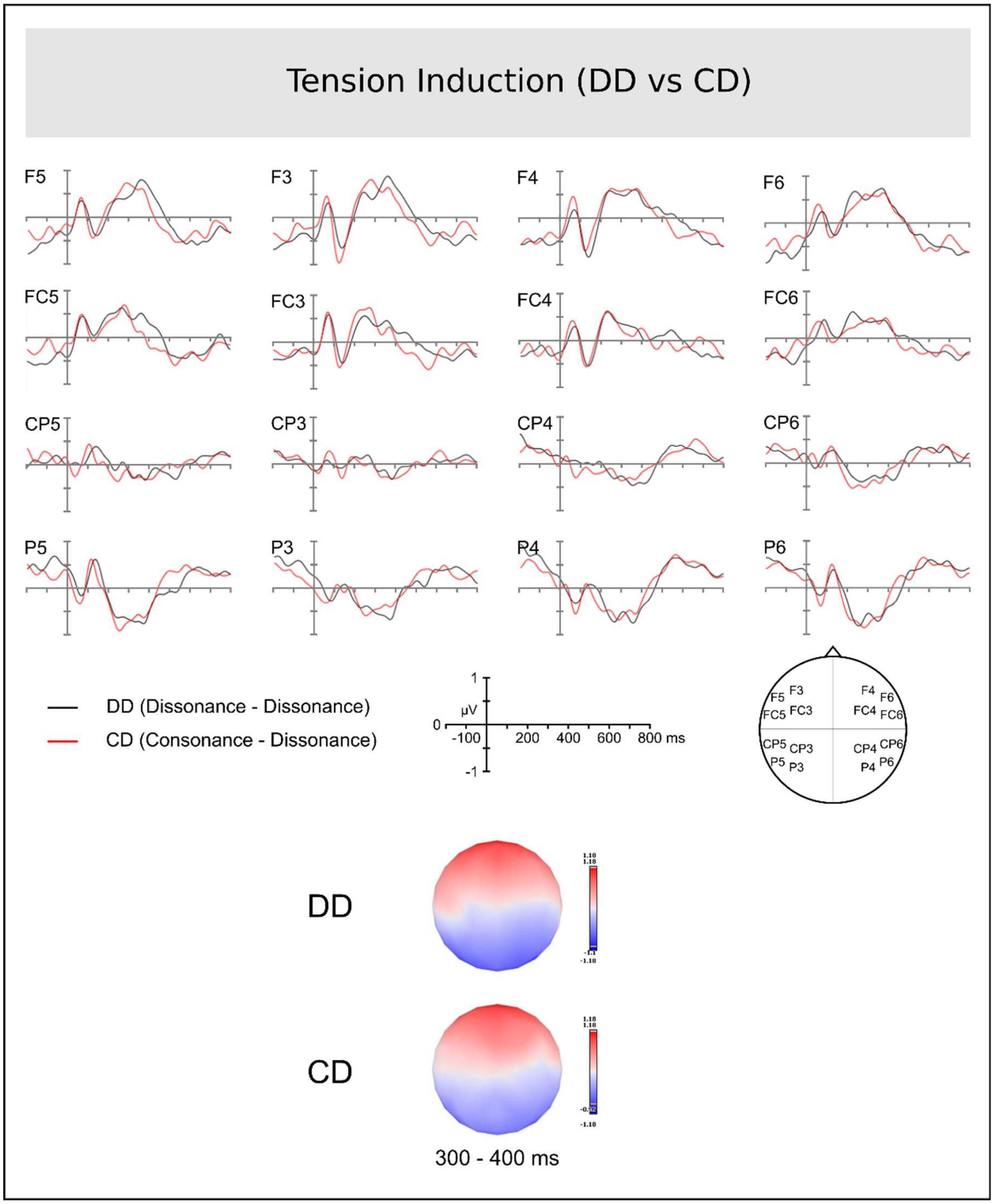
3.4 Discussion

The aim of the present study was to disentangle the effects of tension-resolution patterns from the effects of basic sensory consonance-dissonance. Two-dyad sequences were constructed, manipulating the relative level of consonance-dissonance of intervals to build short tension-resolution patterns.

A consonant interval followed by a dissonant one would create a 'tension induction' effect, vice-versa, a dissonant interval followed by a consonant one would create a 'resolution' effect. Each of these two conditions was compared with a control condition that shared the same type of interval presented second while differed in the type of interval presented first. Brain signals evoked during the second dyad presentation were analyzed so that any emerging difference would result from the type of interval presented previously and not from a difference in the perceived roughness.

Figure 21

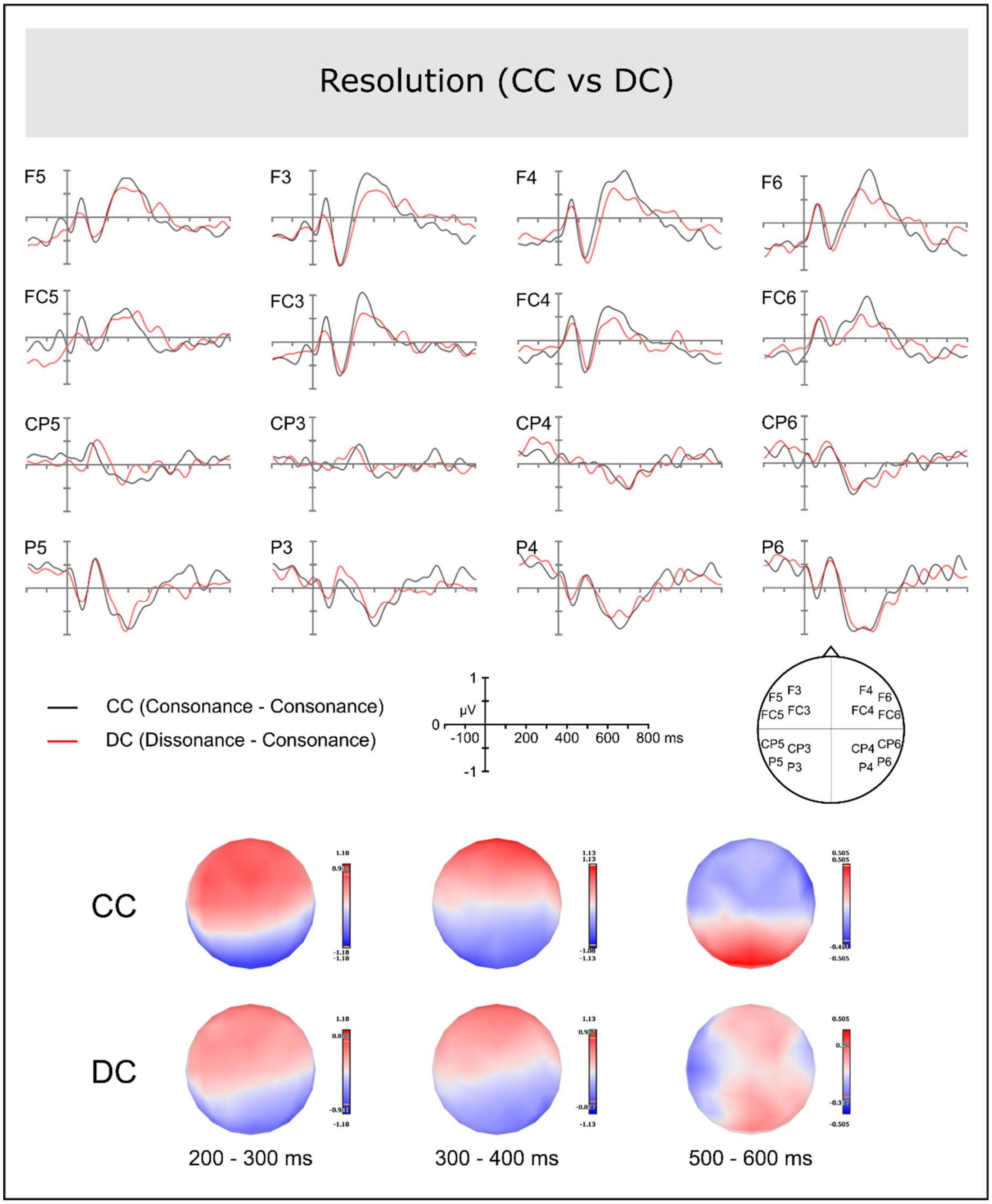
ERPs Comparison Between the Consonance - Dissonance Condition and Dissonance - Dissonance Condition



Note. Grand averaged ERPs in the CD (black) vs. DC (red) condition recorded in four regions: left anterior (F5, F3, FC5, FC3), right anterior (F4, F6, FC4, FC6), left posterior (CP5, CP3, P5, P3), and right posterior (CP4, CP6, P4, P6). Topographic maps of the average amplitude show the significant effects in the 300 – 400 ms time window.

Figure 22

ERPs Comparison Between the Consonance – Dissonance (CD) Condition and Dissonance - Dissonance (DD) Condition



Note. Grand averaged ERPs in the CD (black) vs. DC (red) condition recorded in four regions: left anterior (F5, F3, FC5, FC3), right anterior (F4, F6, FC4, FC6), left posterior (CP5, CP3, P5, P3), and right posterior (CP4, CP6, P4, P6). Topographic maps of the average amplitude show the significant effects within three temporal windows (200 – 300, 300 – 400, and 500 – 600 ms).

A shift in consonance-dissonance from the first interval to the second elicits an overall neural response that is more evident during the resolution of tension than the tension induction, indicating that harmonic resolutions are more salient compared to the harmonic creation of tension. The emerged ERP components indicate that tension-resolution patterns built on harmonic intervals involve not only early processing (P2, P300) related to changes in roughness and attentive processes, but also later processes (LPC) that may be linked to semantic-like processes of meaning attribution (Koelsch, 2011; Steinbeis & Koelsch, 2008).

The P2 component has been found in previous studies involving harmonic intervals perception. Itoh and colleagues (2003) investigated the cortical processes underlying perception of musical consonance by presenting a random sequence of harmonic intervals (minor second, major third, tritone/augmented fourth, perfect fifth, major sixth). The amplitude of a P2 component peaking at 160-180 ms was significantly modulated by the type of interval being most positive for seven semitones (perfect fifth), which had the simplest frequency ratio (2:3). The authors concluded that this component is associated with the processing of consonance. Similar results were found by Kung et al. (2014), who observed a modulation of the P2 component due to the pitch interval or roughness in musicians and nonmusicians, respectively. The positive amplitude was greater for fifths vs. tritone and stimuli with no sensory roughness. Moreover, this component seems to be modulated by musical training as musicians show a larger P2 amplitude in response to complex sound spectra compared to nonmusicians (Shahin et al., 2005). These findings are consistent with the present results since the P2 component was elicited only by consonant intervals. In general, the P2 component seems to be associated with the processing of basic harmonic processing of consonance and modulated by the level of roughness. Notably, Crespo-Bojorque et al. (2018) found a similar positive component peaking around 200 ms during using an oddball paradigm containing consonant and dissonant harmonic intervals. This component was elicited when rare consonant stimuli were presented during a sequence of dissonant stimuli, both in musicians and nonmusicians, while rare dissonant stimuli presented during a sequence of consonant stimuli evoked a response only in musicians. These results could explain why the P2 component was visible during the presentation of consonant stimuli in the 'resolution' condition and not during the presentation of dissonant stimuli in the 'tension induction' condition.

The P300 component has been often linked to attentive processes both in task-relevant and -nonrelevant conditions (Koelsch, 2012; Luck, 2014). In the present study, this component started around 300 ms from the onset and was maximal over right frontal sites resembling a P3b component, thought to reflect context-based memory updating and previously observed in studies on expectancies violations (James et al., 2008, 2017; Koelsch et al., 2007; Polich & Criado, 2006). The P3b component is also thought to underlie also decisional processes during the conscious recognition and detection of a target stimulus (Koelsch, 2012). The present task required the retention of the first stimulus during the second stimulus presentation, an active comparison between the two stimuli, and a decisional step before responding. These processes should have been active during all the conditions. Therefore, a difference in the P3b amplitude between the 'resolution' condition vs. the 'consonance' condition could reflect the detection of specific acoustic properties during the active comparison of stimuli.

Finally, the LPC that emerged in this study is the component whose scalp distribution differs the most between the 'resolution' and 'consonance' condition, with an opposite polarity over the mid-frontal and left parietal regions. Previous studies on the expectancy violation where the final chords were manipulated highlighted components associated with tension induction. For example, harmonically unexpected/incongruous events have been shown to elicit LPCs, peaking around between 500 and 600 ms (Besson & Macar, 1987; Besson & Faïta, 1995; Besson et al., 1994; Janata, 1995; Levett & Martin, 1992; Patel, 1998).

However, since the present study employed short (1600 ms) stimuli sequences designed to create a minimal harmonic context and not evoke expectancies, we can hypothesize that the LPC elicited here is associated with the shifting from a state of perceived instability to stability. In general, this component may reflect structural integration processes and possibly processes of structural repair (Koelsch & Siebel, 2005).

Moreover, the P3b and the LPC could be related somehow, but it remains to be specified whether the LPC is a late P3 and how the processes of structural reanalysis and repair are possibly related to context-updating (Koelsch, 2011). In light of these considerations, it remains unclear why this component did not also emerge in the ‘tension induction’ condition. Although some differences can be noted by visual inspecting the resulting ERP, especially in the mid-frontal and left parietal sites (see Figure 21), it could be that the statistical analysis failed to reveal a significant effect for this component.

Finally, some limitations can be acknowledged regarding the behavioral performance of participants at the tension detection task. Although the perceived tension associated with specific musical intervals in this study was expected to be easily detectable according to the previous experiments' results, a considerable gap in response accuracy emerged between sequences containing stimuli whose sensory consonance-dissonance was relatively more evident compared to those where differences were less evident. Specifically, participants encountered more difficulties when they had to notice differences in perceived tension between perfect fourth, perfect fifth, and augmented fourth (P4, P5, A4) compared to the unison, minor second, and inverted major seventh (P0, m2, iM7) (see Table 14). This difference was larger in the tension induction condition compared to the resolution condition. However, the overall difference in response accuracy between the two conditions was not significant; for this reason, it is unlikely that this difference in the detection task could have influenced ERP results somehow. Future studies involving professional musicians could minimize any possible distortion introduced by the task difficulty and possibly enhance the ERP effects observed in the present study.

General conclusion

Tension is a key component of everyone's emotional experience during music listening. Music elements can modulate the perception of tension at various levels of complexity. At the most basic level, we can find the psychoacoustical features (e.g., pitch, timbre, sensory consonance-dissonance) of discrete elements (i.e., single notes or synchronous groups of notes). The aim of this dissertation was to investigate the role of one of these features, namely, sensory consonance-dissonance, in the perception of tension-resolution. Three empirical studies explored perceived tension in harmonic intervals using a novel proprioceptive device that emphasizes the embodied nature of tension by matching psychological and muscular tension. A fourth study investigated the brain electrical processing of tension-resolution built on consonance-dissonance transitions.

The results of the first study allowed us to find the psychophysical equation for the transformation of physical force into subjective force. Two subsequent behavioral studies showed the relative amount of tension evoked by the different harmonic intervals and its relationship with two other dimensions: valence and perceived movement. Sensory dissonance/roughness and pitch both modulated perceived tension ratings, with more dissonant and high pitch intervals judged as more tense. The inclusion of noises allowed to disentangle tension and valence as two distinct dimensions, although strongly (inversely) related. The comparison between two different rating methods (VAS and proprioceptive device) showed that the proprioceptive device generates more conservative tension judgments. These results should orient future studies on musical tension both for a careful selection of stimuli and rating methods. Further research is needed to explore the embodied nature of perceived tension, not only in terms of physiological activation (arousal) but also in terms of muscular contraction and action tendency.

The last study showed that tension-resolution could also be evoked by very basic musical stimuli (i.e., two-dyads sequences), provided with a minimal harmonic context. A shift in consonance-dissonance is processed by the brain not only at a sensory level (i.e., extraction of harmonic features, consonance-dissonance) but also at a deeper syntactic-like level that could be related to the subjective experience of tension. Future studies could employ ERP to investigate the brain dynamics associated with changes in other psychoacoustical properties known to modulate the subjective experience of tension (e.g., timbre, see Farbood & Price, 2017).

Ultimately, the understanding of the processes underlying the perception of musical tension could inform general theories of emotion and empirical studies that use music as a means to induce emotions. Moreover, the possibility to evoke subjective experiences of tension could find application in the field of health and clinical psychology. For example, music stimuli could be specifically designed to be used in combination with well-known techniques for the treatment of anxiety symptoms, such as the progressive muscle relaxation technique, based on the tension-release of muscle districts (Jacobson, 1938).

References

American National Standards Institute. (2007). *ANSI S34-2007 Procedure for the computation of loudness of steady sounds*.

Aures, W. (1985). Ein berechnungsverfahren der rauhigkeit. *Acta Acustica united with Acustica*, *58*(5), 268-281.

Berger, H. (1929). Über das Elektrenkephalogramm des Menschen. *Archiv Für Psychiatrie Und Nervenkrankheiten*. https://doi.org/10.1007/BF01797193

Besson, M., & Macar, F. (1987). An event-related potential analysis of incongruity in music and other non-linguistic contexts. *Psychophysiology*, *24*(1), 14-25. https://doi.org/10.1111/j.1469-8986.1987.tb01853.x

Besson, M., & Faïta, F. (1995). An event-related potential (ERP) study of musical expectancy: Comparison of musicians with nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(6), 1278. https://doi.org/10.1037/0096-1523.21.6.1278

Besson, M., Faïta, F., & Requin, J. (1994). Brain waves associated with musical incongruities differ for musicians and non-musicians. *Neuroscience letters*, *168*(1-2), 101-105. https://doi.org/10.1016/0304-3940(94)90426-X

Bidelman, G. M., & Grall, J. (2014). Functional organization for musical consonance and tonal pitch hierarchy in human auditory cortex. *NeuroImage*, *101*, 204–214. https://doi.org/10.1016/j.neuroimage.2014.07.005

Bidelman, G. M., & Krishnan, A. (2009). Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. *Journal of Neuroscience*, *29*(42), 13165-13171. https://doi.org/10.1523/JNEUROSCI.3900-09.2009

Bigand, E., & Parncutt, R. (1999). Perceiving musical tension in long chord sequences. *Psychological Research*, *62*(4), 237–254. https://doi.org/10.1007/s004260050053

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 & Psychophysics*, *58*(1), 125–141. https://doi.org/10.3758/BF03205482

Bonetti, L., & Costa, M. (2018). Pitch-verticality and pitch-size cross-modal interactions. *Psychology of Music*, *46*(3), 340–356. https://doi.org/10.1177/0305735617710734

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial vision*, *10*(4), 433-436. https://doi.org/10.1163/156856897X00357

Brattico, E., Tervaniemi, M., Näätänen, R., & Peretz, I. (2006). Musical scale properties are automatically processed in the human auditory cortex. *Brain research*, *1117*(1), 162-174. *Brain Research*. https://doi.org/10.1016/j.brainres.2006.08.023

Brunet, D., Murray, M. M., & Michel, C. M. (2011). Spatiotemporal analysis of multichannel EEG: CARTOOL. *Computational Intelligence and Neuroscience*, *2011*. https://doi.org/10.1155/2011/813870

Burnsed, V., & Sochinski, J. (1998). The relationship between dynamics and tension in Haydn's symphony 104: A developmental study. *Psychomusicology: A Journal of Research in Music Cognition*, *17*(1-2), 19–35. https://doi.org/10.1037/h0094061

Chapados, C., & Levitin, D. J. (2008). Cross-modal interactions in the experience of musical performances: Physiological correlates. *Cognition*, *108*(3), 639–651. https://doi.org/10.1016/j.cognition.2008.05.008

Choppin, S., Trost, W., Dondaine, T., Millet, B., Drapier, D., Vérin, M., ... & Grandjean, D. (2016). Alteration of complex negative emotions induced by music in euthymic patients with bipolar disorder. *Journal of affective disorders*, *191*, 15-23. https://doi.org/10.1016/j.jad.2015.10.063

Costa, M., Fine, P., & Ricci Bitti, P. E. (2004). Interval distributions, mode, and tonal strength of melodies as predictors of perceived emotion. *Music Perception*, *22*(1), 1-14. https://doi.org/10.1525/mp.2004.22.1.1

Costa, M., Ricci Bitti, P. E., & Bonfiglioli, L. (2000). Psychological connotations of harmonic musical intervals. *Psychology of Music*, *28*(1), 4–22. https://doi.org/10.1177/0305735600281002

Crespo-Bojorque, P., Monte-Ordoño, J., & Toro, J. M. (2018). Early neural responses underlie advantages for consonance over dissonance. *Neuropsychologia*, *117*, 188-198. https://doi.org/10.1016/j.neuropsychologia.2018.06.005

Daniel, P., & Weber, R. (1997). Daniel, P., & Weber, R. (1997). Psychoacoustical roughness: Implementation of an optimized model. *Acta Acustica united with Acustica*, *83*(1), 113-123.

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

Evans, K. K., & Treisman, A. (2011). Natural cross-modal mappings between visual and auditory features. *Journal of vision*, *10*(1), 6–6. https://doi.org/10.1167/10.1.6

Evans, P., & Schubert, E. (2008). Relationships between expressed and felt emotions in music. *Musicae Scientiae*, *12*(1), 75–99. https://doi.org/10.1177/102986490801200105

Fallow, P. M., Gardner, J. L., & Magrath, R. D. (2011). Sound familiar? Acoustic similarity provokes responses to unfamiliar heterospecific alarm calls. *Behavioral Ecology*, *22*(2), 401–410. https://doi.org/10.1093/beheco/arq221

Fang, L., Shang, J., & Chen, N. (2017). Perception of western musical modes: a Chinese study. *Frontiers in Psychology*, *8*, 1905. https://doi.org/10.3389/fpsyg.2017.01905

Farbood, M. M. (2012). A parametric, temporal model of musical tension. *Music Perception*, *29*(4), 387–428. https://doi.org/10.1525/mp.2012.29.4.387

Farbood, M. M., & Price, K. C. (2017). The contribution of timbre attributes to musical tension. *The Journal of the Acoustical Society of America*, *141*(1), 419–427. https://doi.org/10.1121/1.4973568

Farbood, M. M., & Upham, F. (2013). Interpreting expressive performance through listener judgments of musical tension. *Frontiers in psychology*, *4*, 998. https://doi.org/10.3389/fpsyg.2013.00998

Fredrickson, W. E. (1995). A comparison of perceived musical tension and aesthetic response. *Psychology of Music*, *23*(1), 81-87. https://doi.org/10.1177/0305735695231006

Fredrickson, W. E. (1997). Elementary, middle, and high school student perceptions of tension in music. *Journal of Research in Music Education*, *45*(4), 626–635. https://doi.org/10.2307/3345427

Fredrickson, W. E. (1999). Effect of musical performance on perception of tension in Gustav Hoist's First Suite in E-flat. *Journal of Research in Music Education*, *47*(1), 44-52. https://doi.org/10.2307/3345827

Fredrickson, W. E. (2000). Perception of tension in music: Musicians versus nonmusicians. *Journal of music Therapy*, *37*(1), 40-50. https://doi.org/10.1093/jmt/37.1.40

Fredrickson, W. E., & Coggiola, J. C. (2003). A comparison of music majors' and nonmajors' perceptions of tension for two selections of jazz music. *Journal of Research in Music Education*, *51*(3), 259–270. https://doi.org/10.2307/3345378

Fredrickson, W. E., & Johnson, C. M. (1996). The effect of performer use of rubato on listener perception of tension in Mozart. *Psychomusicology: A Journal of Research in Music*

Cognition, *15*(1–2), 78–86. https://doi.org/10.1037/h0094078

Frego, R. J. D., & Frego R.J. David, R. J. D. (1999). Effects of aural and visual conditions on response to perceived artistic tension in music and dance. *Journal of Research in Music Education*, *47*(1), 31–43. https://doi.org/10.2307/3345826

Genesis, S. A. (2009). Loudness Toolbox. For Matlab.

Gingras, B., Marin, M. M., Puig-Waldmüller, E., & Fitch, W. T. (2015). The eye is listening: Music-induced arousal and individual differences predict pupillary responses. *Frontiers in human neuroscience*, *9*, 619. https://doi.org/10.3389/fnhum.2015.00619

Gingras, B., Pearce, M. T., Goodchild, M., Dean, R. T., Wiggins, G., & McAdams, S. (2016). Linking melodic expectation to expressive performance timing and perceived musical tension. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(4), 594–609. https://doi.org/10.1037/xhp0000141

Goodchild, M., Gingras, B., & McAdams, S. (2016). Analysis, performance, and tension perception of an unmeasured prelude for harpsichord. *Music Perception: An Interdisciplinary Journal*, *34*(1), 1-20. https://doi.org/10.1525/mp.2016.34.1.1

Granot, R. Y., & Eitan, Z. (2011). Musical tension and the interaction of dynamic auditory parameters. *Music Perception*, *28*(3), 219-246. https://doi.org/10.1525/mp.2011.28.3.219

Hackworth, R. S., & Fredrickson, W. E. (2010). The effect of text translation on perceived musical tension in Debussy's noel des enfants qui n'ont plus de maisons. *Journal of Research in Music Education*, *58*(2), 184-195. https://doi.org/10.1177/0022429410369835

Helmholtz, H. von (1877). Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik (6th ed.). Braunschweig: Vieweg, 1913; trans. by A. J. Ellis as On the sensations of tone as a physiological basis for the theory of music (1885). Reprinted (1954). New York: Dover.

Huron, D. (2008). *Sweet anticipation: Music and the psychology of expectation*. MIT press.

Hutchinson, W., & Knopoff, L. (1978). The acoustic component of western consonance. *Interface*, *7*(1), 1–29. https://doi.org/10.1080/09298217808570246

Ilie, G., & Thompson, W. F. (2006). A comparison of acoustic cues in music and speech for three dimensions of affect. *Music Perception*, *23*(4), 319-330. https://doi.org/10.1525/mp.2006.23.4.319

Itoh, K., Suwazono, S., & Nakada, T. (2003). Cortical processing of musical consonance: an evoked potential study. *NeuroReport*, *14*(18), 2303–2306. https://doi.org/10.1097/00001756-200312190-00003

Itoh, K., Suwazono, S., & Nakada, T. (2011). Central auditory processing of noncontextual consonance in music: an evoked potential study. *The Journal of the Acoustical Society of America*, *128*(6), 3781–3787. https://doi.org/10.1121/1.3500685

Iwanaga, M., Ikeda, M., & Iwaki, T. (1996). The effects of repetitive exposure to music on subjective and physiological responses. *Journal of Music Therapy*, *33*(3), 219-230. https://doi.org/10.1093/jmt/33.3.219

Iwanaga, M., Kobayashi, A., & Kawasaki, C. (2005). Heart rate variability with repetitive exposure to music. *Biological Psychology*, *70*(1), 61–66. https://doi.org/10.1016/j.biopsycho.2004.11.015

Iwanaga, M., & Moroki, Y. (1999). Subjective and physiological responses to music stimuli controlled over activity and preference. *Journal of Music Therapy*, *36*(1), 26–38. https://doi.org/10.1093/jmt/36.1.26

Jacobson, E. (1938). Progressive muscle relaxation. J Abnorm Psychol, 75(1), 18.

James, C. E., Britz, J., Vuilleumier, P., Hauert, C. A., & Michel, C. M. (2008). Early neuronal responses in right limbic structures mediate harmony incongruity processing in musical experts. *NeuroImage*, *42*(4), 1597–1608. https://doi.org/10.1016/j.neuroimage.2008.06.025

James, C. E., Cereghetti, D. M., Roullet Tribes, E., & Oechslin, M. S. (2015). Electrophysiological evidence for a specific neural correlate of musical violation expectation in primary-school children. *NeuroImage*, *104*, 386–397. https://doi.org/10.1016/j.neuroimage.2014.09.047

James, C. E., Oechslin, M. S., Michel, C. M., & De Pretto, M. (2017). Electrical neuroimaging of music processing reveals mid-latency changes with level of musical expertise. *Frontiers in neuroscience*, *11*, 613. https://doi.org/10.3389/fnins.2017.00613

Janata, P. (1995). ERP measures assay the degree of expectancy violation of harmonic contexts in music. *Journal of Cognitive Neuroscience*, *7*(2), 153-164. https://doi.org/10.1162/jocn.1995.7.2.153

Jenni, R., Oechslin, M. S., & James, C. E. (2017). Impact of major and minor mode on EEG frequency range activities of music processing as a function of expertise. *Neuroscience Letters*, *647*, 159–164. https://doi.org/10.1016/j.neulet.2017.03.022

Kallinen, K., & Ravaja, N. (2006). Emotion perceived and emotion felt: Same and different. *Musicae Scientiae*, *10*(2), 191-213. https://doi.org/10.1177/102986490601000203

Kleinsmith, A. L., Friedman, R. S., & Neill, W. T. (2017). Reinvestigating the impact of final ritardandi on music evaluation and felt musical tension. *Psychomusicology: Music, Mind, and Brain*, *27*(4), 267–280. https://doi.org/10.1037/pmu0000195

Koelsch, S. (2011). Toward a neural basis of music perception–a review and updated model. *Frontiers in psychology*, *2*, 110. https://doi.org/10.3389/fpsyg.2011.00110

Koelsch, S. (2012). *Brain and music*. Hoboken, NJ: Wiley-Blackwell.

Koelsch, S., Gunter, T., Friederici, A. D., & Schröger, E. (2000). Brain indices of music processing:“nonmusicians” are musical. *Journal of cognitive neuroscience*, *12*(3), 520-541. https://doi.org/10.1162/089892900562183

Koelsch, S., Jentschke, S., Sammler, D., & Mietchen, D. (2007). Untangling syntactic and sensory processing: An ERP study of music perception. *Psychophysiology*, *44*(3), 476–490. https://doi.org/10.1111/j.1469-8986.2007.00517.x

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Friederici, A. D. (2004). Music, language and meaning: Brain signatures of semantic processing. *Nature Neuroscience*, *7*(3), 302–307. https://doi.org/10.1038/nn1197

Koelsch, S., Rohrmeier, M., Torrecuso, R., & Jentschke, S. (2013). Processing of hierarchical syntactic structure in music. *Proceedings of the National Academy of Sciences*, *110*(38), 15443–15448. https://doi.org/10.1073/pnas.1300272110

Koelsch, S., & Siebel, W. A. (2005). Towards a neural basis of music perception. *Trends in Cognitive Sciences*, *9*(12), 578–584. https://doi.org/10.1016/j.tics.2005.10.001

Krumhansl, C. L. (1996). A perceptual analysis of Mozart's Piano Sonata K. 282: Segmentation, tension, and musical ideas. *Music perception*, *13*(3), 401-432. https://doi.org/10.2307/40286177

Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 51*(4), 336–353. https://doi.org/10.1037/1196-1961.51.4.336

Krumhansl, C. L. (2002). Music: A link between cognition and emotion. *Current directions in psychological science*, *11*(2), 45-50. https://doi.org/10.1111/1467-8721.00165

Kung, C. C., Hsieh, T. H., Liou, J. Y., Lin, K. J., Shaw, F. Z., & Liang, S. F. (2014). Musicians and non-musicians’ different reliance of features in consonance perception: A behavioral and ERP study. *Clinical Neurophysiology*, *125*(5), 971-978. https://doi.org/10.1016/j.clinph.2013.10.016

Lahdelma, I., & Eerola, T. (2016a). Mild dissonance preferred over consonance in single chord perception. *i-Perception, 7*(3), Article 2041669516655812. https://doi.org/10.1177/2041669516655812

Lahdelma, I., & Eerola, T. (2016b). Single chords convey distinct emotional qualities to both naïve and expert listeners. *Psychology of Music*, *44*(1), 37–54. https://doi.org/10.1177/0305735614552006

Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, 963-974. https://doi.org/10.2307/2529876

Lartillot, O., Toiviainen, P., & Eerola, T. (2008). A Matlab toolbox for music information retrieval. In C. Preisach, H. Burkhardt, L. Schmidt-Thieme, & R. Decker (Eds.), *Data analysis, machine learning, and applications: Studies in classification, data analysis, and knowledge organization* (pp. 261–268). Berlin, Germany: Springer.

Lehne, M., & Koelsch, S. (2015). Toward a general psychological model of tension and suspense. *Frontiers in Psychology*, *6*, 79. https://doi.org/10.3389/fpsyg.2015.00079

Lehne, M., Rohrmeier, M., Gollmann, D., & Koelsch, S. (2012). The influence of different structural features on felt musical tension in two piano pieces by Mozart and Mendelssohn. *Music Perception: An Interdisciplinary Journal*, *31*(2), 171-185. https://doi.org/10.1525/mp.2013.31.2.171

Lehne, M., Rohrmeier, M., & Koelsch, S. (2013). Tension-related activity in the orbitofrontal cortex and amygdala: an fMRI study with music. *Social cognitive and affective neuroscience*, *9*(10), 1515–1523. https://doi.org/10.1093/scan/nst141

Leman, M. (2000, December). Visualization and calculation of the roughness of acoustical musical signals using the synchronization index model (SIM). In *Proceedings of the COST G-6 conference on digital audio effects (DAFX-00)* (pp. 1-6). DAFX, Verona, Italy.

Lerdahl, F. (1996). Calculating tonal tension. *Music Perception*, *13*(3), 319–363. https://doi.org/10.2307/40286174

Lerdahl, F. (2001). *Tonal pitch space*. New York: Oxford University Press.

Lerdahl, F., & Krumhansl, C. L. (2007). Modeling tonal tension. *Music perception*, *24*(4), 329-366. https://doi.org/10.1525/mp.2007.24.4.329

Levett, C., & Martin, F. (1992). The relationship between complex music stimuli and the late components of the event-related potential. *Psychomusicology: A Journal of Research in Music Cognition*, *11*(2), 125–140. https://doi.org/10.1037/h0094126

Link, S., Poulton, E. C., Link, S., & Poulton, E. C. (1991). Bias in quantifying judgments. *The American Journal of Psychology*, *104*(2), 295–301. https://doi.org/10.2307/1423161

Luck, S. J. (2014). *An introduction to the event-related potential technique* (2nd ed.). Cambridge, MA: MIT Press.

Lychner, J. A. (1998). An empirical study concerning terminology relating to aesthetic response to music. *Journal of Research in Music Education*, *46*(2), 303-319. https://doi.org/10.2307/3345630

Madsen, C. K., & Fredrickson, W. E. (1993). The experience of musical tension: a replication of nielsen's research using the continuous response digital interface. *Journal of Music Therapy*, *30*(1), 46–63. https://doi.org/10.1093/jmt/30.1.46

Malmberg, C. F. (1918). The perception of consonance and dissonance. *Psychological Monographs, 25*(2), 93–133. https://doi.org/10.1037/h0093119

M Margulis, E. H. (2007). Silences in music are musical not silent: an exploratory study of context effects on the experience of musical pauses. *Music Perception*, *24*(5), 485-506. https://doi.org/10.1525/mp.2007.24.5.485

McAdams, S., Douglas, C., & Vempala, N. N. (2017). Perception and modeling of affective qualities of musical instrument sounds across pitch registers. *Frontiers in Psychology*, *8*, 153. https://doi.org/10.3389/fpsyg.2017.00153

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

Misenhelter, D. (2001). An investigation of music and nonmusic majors' responses to musical tension and dynamics in Beethoven's Symphony No. 7. *Missouri Journal of Research in Music Education*, *38*, 56-67.

Moore, B. C., & 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. https://doi.org/10.1121/1.389861

Nielsen, F. V. (1983). *Oplevelse afmusikalsk spaending* [The experience ofmusical tension]. Copenhagen: Akademisk Forlag.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*(1), 97-113. https://doi.org/10.1016/0028-3932(71)90067-4

Pagès-Portabella, C., & Toro, J. M. (2020). Dissonant endings of chord progressions elicit a larger

ERAN than ambiguous endings in musicians. *Psychophysiology*, *57*(2), e13476. https://doi.org/10.1111/psyp.13476

Paller, K. A., McCarthy, G., & Wood, C. C. (1992). Event-related potentials elicited by deviant endings to melodies. *Psychophysiology*, *29*(2), 202-206. https://doi.org/10.1111/j.1469-8986.1992.tb01686.x

Parncutt, R. (2012). *Harmony: A psychoacoustical approach* (Vol. 19). Berlin, Germany: Springer-Verlag.

Patel, A. D. (1998). Syntactic processing in language and music: different cognitive operations, similar neural resources?. *Music perception*, *16*(1), 27-42. https://doi.org/10.2307/40285775

Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and clinical neurophysiology*, *72*(2), 184-187. https://doi.org/10.1016/0013-4694(89)90180-6

Pinheiro, J., & Bates, D. (2006). *Mixed-effects models in S and S-PLUS*. Berlin, Germany: Springer-Verlag.

Plomp, R., & Levelt, W. J. M. (1965). Tonal consonance and critical bandwidth. *The journal of the Acoustical Society of America*, *38*(4), 548-560. https://doi.org/https://doi.org/10.1121/1.1909741

Plomp, R., & Steeneken, H. J. M. (1968). Interference between two simple tones. *The Journal of the Acoustical Society of America*, *43*(4), 883-884. https://doi.org/10.1121/1.1910916

Polich, J., & Criado, J. R. (2006). Neuropsychology and neuropharmacology of P3a and P3b. *International Journal of Psychophysiology*, *60*(2), 172-185. https://doi.org/10.1016/j.ijpsycho.2005.12.012

Pressnitzer, D., McAdams, S., Winsberg, S., & Fineberg, J. (2000). Perception of musical tension for nontonal orchestral timbres and its relation to psychoacoustic roughness. *Perception and Psychophysics*, *62*(1), 66–80. https://doi.org/10.3758/BF03212061

Proverbio, A. M., Orlandi, A., & Pisanu, F. (2016). Brain processing of consonance/dissonance in musicians and controls: a hemispheric asymmetry revisited. *European Journal of Neuroscience*, *44*(6), 2340-2356. https://doi.org/10.1111/ejn.13330

Regnault, P., Bigand, E., & Besson, M. (2001). Different brain mechanisms mediate sensitivity to sensory consonance and harmonic context: Evidence from auditory event-related brain potentials. *Journal of Cognitive Neuroscience*, *13*(2), 241–255. https://doi.org/10.1162/089892901564298

Rozin, A., Rozin, P., & Goldberg, E. (2004). The feeling of music past: How listeners remember musical affect. *Music Perception*, *22*(1), 15-39. https://doi.org/10.1525/mp.2004.22.1.15

Rusconi, E., Kwan, B., Giordano, B. L., Umiltà, C., & Butterworth, B. (2006). Spatial representation of pitch height: the SMARC effect. *Cognition*, *99*(2), 113–129. https://doi.org/10.1016/j.cognition.2005.01.004

Schedl, M., Gómez, E., Trent, E. S., Tkalčič, M., Eghbal-Zadeh, H., & Martorell, A. (2017). On the Interrelation between listener characteristics and the perception of emotions in classical orchestra music. *IEEE Transactions on Affective Computing*, *9*(4), 507-525. https://doi.org/10.1109/TAFFC.2017.2663421

Schenker, H. (1935). *Free composition*. Tr. E. Oster. New York, Longman, 1979.

Schimmack, U., & Grob, A. (2000). Dimensional models of core affect: A quantitative comparison by means of structural equation modeling. *European Journal of Personality*, *14*(4), 325-345. https://doi.org/10.1002/1099-0984(200007/08)14:4<325::AID-PER380>3.0.CO;2-I

Schimmack, U., & Rainer, R. (2002). Experiencing activation: Energetic arousal and tense arousal are not mixtures of valence and activation. *Emotion*, *2*(4), 412–417. https://doi.org/10.1037/1528-3542.2.4.412

Schön, D., Regnault, P., Ystad, S., & Besson, M. (2005). Sensory consonance: An ERP study. *Music Perception*, *23*(2), 105–118. https://doi.org/10.1525/mp.2005.23.2.105

Schuetze, P., & Zeskind, P. S. (2001). Relations between women's depressive symptoms and perceptions of infant distress signals varying in pitch. *Infancy*, *2*(4), 483-499. https://doi.org/10.1177/153270802237357

Sethares, W. A. (2005). *Tuning, Timbre, Spectrum, Scale* (2nd ed.). London, UK: Springer.

Shahin, A., Roberts, L. E., Pantev, C., Trainor, L. J., & Ross, B. (2005). Modulation of P2 auditory-evoked responses by the spectral complexity of musical sounds. *Neuroreport*, *16*(16), 1781-1785. https://doi.org/10.1097/01.wnr.0000185017.29316.63

Smith, N. A., & Cuddy, L. (2003). Perceptions of musical dimensions in Beethoven's Waldsiein sonata: An application of Tonal Pitch Space theory. *Musicae Scientiae*, *7*(1), 7-34. https://doi.org/10.1177/102986490300700102

Sobin, C., & Alpert, M. (1999). Emotion in speech: The acoustic attributes of fear, anger, sadness, and joy. *Journal of Psycholinguistic Research*, *28*(4), 347–365. https://doi.org/10.1023/A:1023237014909

Soltis, J. (2004). The signal functions of early infant crying. *Behavioral and Brain Sciences*, *27*(4), 443–458. https://doi.org/10.1017/S0140525X0400010X

Steinbeis, N., & Koelsch, S. (2008). Shared neural resources between music and language indicate semantic processing of musical tension-resolution patterns. *Cerebral Cortex*, *18*(5), 1169-1178. https://doi.org/10.1093/cercor/bhm149

Steinbeis, N., Koelsch, S., & Sloboda, J. A. (2006). The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. *Journal of cognitive neuroscience*, *18*(8), 1380-1393. https://doi.org/10.1162/jocn.2006.18.8.1380

Stevens, J. C. (1989). Static and dynamic exertion: A psychophysical similarity and dissimilarity. In *Psychophysics in action* (pp. 81-93). Springer, Berlin, Heidelberg.

Stevens, S. S. (1958). Problems and methods of psychophysics. *Psychological Bulletin*, *55*(4), 177–196. https://doi.org/10.1037/h0044251

Stevens, S. S. (1959). Cross-modality validation of subjective scales for loudness, vibration, and electric shock. *Journal of Experimental Psychology*, *57*(4), 201–209. https://doi.org/10.1037/h0048957

Stevens, S. S. (1960). The psychophysics of sensory function. *American Scientist*, *48*(2), 226–253.

Stevens, S. S. (1975). *Psychophysics: introduction to its perceptual, neural, and social prospects* . Hoboken, NJ: Wiley.

Sturm, I., Dähne, S., Blankertz, B., & Curio, G. (2015). Multi-variate EEG analysis as a novel tool to examine brain responses to naturalistic music stimuli. *PloS one*, *10*(10), e0141281. https://doi.org/10.1371/journal.pone.0141281

Sun, L., Feng, C., & Yang, Y. (2020). Tension experience induced by nested structures in music. *Frontiers in Human Neuroscience*, *14, 210*. https://doi.org/10.3389/fnhum.2020.00210

Susini, P., & McAdams, S. (2000). Psychophysical validation of a proprioceptive device by cross-modal matching of loudness. *Acta Acustica United with Acustica*, *86*(June), 515–525.

Susini, P., Mcadams, S., & Smith, B. K. (2002). Global and continuous loudness estimation of time-varying levels. *Acta Acustica United with Acustica*, *88*(4), 536–548. https://doi.org/10.1016/j.jss.2015.05.022

Terhardt, E. (1978). Psychoacoustic evaluation of musical sounds. *Perception & Psychophysics*, *23*(6), 483-492.

Thayer, R. (1989). *The biopsychology of mood and arousal*. Oxford University Press.

Trolio, M. (1976). Theories of affective response to music. *Contributions to Music Education*, *4*, 1–20.

Van Doren, C. L. (1996). Halving and doubling isometric force: Evidence for a decelerating psychophysical function consistent with an equilibrium-point model of motor control. *Perception & psychophysics*, *58*(4), 636-647.

Vassilakis, P. N. (2001). Perceptual and physical properties of amplitude fluctuation and their musical significance. *Doctoral dissertation*. University of California, Los Angeles.

Verleger, R. (1990). P3-evoking wrong notes: unexpected, awaited, or arousing?. *International Journal of Neuroscience*, *55*(2-4), 171-179. https://doi.org/10.3109/00207459008985972

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

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., & Levitin, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition*, *101*(1), 80–113. https://doi.org/10.1016/j.cognition.2005.09.003

Vines, B. W., Nuzzo, R. L., & Levitin, D. J. (2005). Analyzing temporal dynamics in music: differential calculus, physics, and functional data analysis techniques. *Music Perception*, *23*(2), 137–152. https://doi.org/10.1525/mp.2005.23.2.137

Virtala, P., Berg, V., Kivioja, M., Purhonen, J., Salmenkivi, M., Paavilainen, P., & Tervaniemi, M. (2011). The preattentive processing of major vs. minor chords in the human brain: An event-related potential study. *Neuroscience Letters*, *487*(3), 406–410. https://doi.org/10.1016/j.neulet.2010.10.066

Virtala, P., Huotilainen, M., Partanen, E., Fellman, V., & Tervaniemi, M. (2013). Newborn infants' auditory system is sensitive to Western music chord categories. *Frontiers in psychology*, *4*, 492. https://doi.org/10.3389/fpsyg.2013.00492

Williams, L. R., Fredrickson, W. E., & Atkinson, S. (2011). Focus of attention to melody or harmony and perception of music tension: An exploratory study. *International Journal of Music Education*, *29*(1), 72–81. https://doi.org/10.1177/0255761410372725

Wundt, W. (1896). *Grundriss der Psychologie [Outline of psychology]*. Leipzig: Engelmann.

Wundt, W. (1911). *Grundzüge der Physiologischen Psychologie [Principles of physiological psychology]*. Leipzig: Engelmann.

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: characterization, classification, and measurement. *Emotion*, *8*(4), 494–521. https://doi.org/10.1037/1528-3542.8.4.494

Zeskind, P. S., & Marshall, T. R. (1988). The relation between variations in pitch and maternal perceptions of infant crying. *Child Development*, 193-196. https://doi.org/10.2307/1130401

Zhou, L., Liu, F., Jiang, J., Jiang, H., & Jiang, C. (2019). Abnormal neural responses to harmonic syntactic structures in congenital amusia. *Psychophysiology*, *56*(9), e13394. https://doi.org/10.1111/psyp.13394

Zwicker, E., & Fastl, H. (1990). *Psychoacoustics: Facts and models*. Berlin, Germany: Springer-Verlag.