Musical Expertise Facilitates Dissonance Detection On Behavioral, Not On Early Sensory Level

Tanja Linnavalli
University of Helsinki, Helsinki, Finland

Juha Ojala
Sibelius Academy of the University of the Arts, Helsinki, Finland

Laura Haveri
University of Helsinki, Helsinki, Finland

Vesa Putkinen
University of Turku, Turku, Finland

Kaisamari Kostilainen, Sirke Seppänen, & Mari Tervaniemi
University of Helsinki, Helsinki, Finland

Consonance and dissonance are basic phenomena in the perception of chords that can be discriminated very early in sensory processing. Musical expertise has been shown to facilitate neural processing of various musical stimuli, but it is unclear whether this applies to detecting consonance and dissonance. Our study aimed to determine if sensitivity to increasing levels of dissonance differs between musicians and non-musicians, using a combination of neural (electroencephalographic mismatch negativity, MMN) and behavioral measurements (conscious discrimination). Furthermore, we wanted to see if focusing attention to the sounds modulated the neural processing. We used chords comprised of either highly consonant or highly dissonant intervals and further manipulated the degree of dissonance to create two levels of dissonant chords. Both groups discriminated dissonant chords from consonant ones neurally and behaviorally. The magnitude of the MMN differed only marginally between the more dissonant and the less dissonant chords. The musicians outperformed the nonmusicians in the behavioral task. As the dissonant chords elicited MMN responses for both groups, sensory dissonance seems to be discriminated in an early sensory level, irrespective of musical expertise, and the facilitating effects of musicianship for this discrimination may arise in later stages of auditory processing, appearing only in the behavioral auditory task.

Sensory consonance and dissonance are essential building blocks of Western music. Studies have shown that discrimination of different chord types occurs at the early sensory level in adults (Brattico et al., 2009), schoolchildren (Virtala, Huotilainen, Putkinen, Makkonen, & Tervaniemi, 2012), and newborn babies (Virtala, Huotilainen, Partanen, Fellman, & Tervaniemi, 2013). It is less clear how the auditory system processes different degrees of dissonance, and how the discrimination of these degrees is manifested neurally and behaviorally. Musicians have been shown to outperform nonmusicians in behaviorally discriminating chords and dissonance from consonance (see, e.g., Kung et al., 2014; Sares, Foster, Allen, & Hyde, 2018; Virtala, Huotilainen, Partanen, & Tervaniemi, 2014), but it is not altogether clear at what stage in the auditory processing stream musical expertise has an effect on consonance/dissonance discrimination.

CONSONANCE AND DISSONANCE

Harmony, an essential element of Western music, relies on the construction of individual and concurrent tones, intervals, and chords, and their successions (e.g., chord progressions). Western *tonal* music is characterized by *functional harmony* in which the tones, intervals, and chords have certain roles in creating music (Kopp, 1995; Rehding, 2003). The functionality of harmony is based to a large extent on how the harmonic structures are perceived as *consonant versus dissonant* (e.g., Rehding, 2019, p. 440). While the context of tonal, functional harmony has been central to, and prevalent in, the scientific study of consonance and dissonance, the phenomenon is not limited to Western tonality. It extends to Western musical practices before and after the so-called Western common-practice tonality, and can be extended to the study of non-Western musical traditions, non-musical sounds, and non-human subjects as well (Dibben, 1999; McDermott, Schultz, Undurraga, & Godoy, 2016; Tenney, 1988; Zarlino, 1558/1982;

versus Bowling, Hoeschele, Gill, & Fitch, 2017; Parncutt & Hair, 2011; Virtala & Tervaniemi, 2017).

Despite their long history, the notions of consonance and dissonance have evaded clear consensus and exact definitions, proving challenging to study empirically (Cazden, 1980; Rehding, 2019, p. 437). There are objective acoustical properties (such as simple integer ratios, or their limited approximations, of fundamental frequencies or other partials of tones) and physiological features (such as those of the critical bandwidths of the basilar membrane) that affect judgments of consonance and dissonance (Bowling & Purves, 2015; Helmholtz, 1877/1954). However, these judgments seem to also be affected by the familiarity, context, cultural background, and musical expertise of the subject (Lahdelma & Eerola, 2016; Parncutt & Hair, 2011; Popescu et al., 2019; Rehding, 2019; Virtala & Tervaniemi, 2017). Furthermore, behavioral research on consonance/dissonance judgments is usually based on studying subjective, experiential, and preference-like ratings using guiding terms such as "pleasantness," "stability," "smoothness," "compatibility" or "relaxation" versus "unpleasantness," "instability," "roughness," "incompatibility" or "tension." This practice is well motivated since the terms and concepts of "consonance" and "dissonance" might not be well understood by nonmusicians or even explicitly by musicians, and the guiding terms generally correlate with the music-theoretical descriptions and estimations of consonance/dissonance (see, e.g., Kuusi, 2001; Lahdelma & Eerola, 2016; Parncutt & Hair, 2011).

To bridge the gap between physical, physiological, psychological, and cultural aspects of consonance/dissonance studies, Parncutt and Hair (2011) proposed a holistic "conceptual structure for Western consonance/dissonance." According to them, the consonance/dissonance of individual harmonic structures (as opposed to successive ones) has "two natural components, smoothness and harmonicity (fusion)" and "the cultural component," that is, familiarity with music one is exposed to, as well as speech and sounds of the environment (see also Harrison & Pearce, 2020; Parncutt & Hair, 2018). Smoothness, roughness, and harmonicity are considered "natural components" because of "their apparent perceptual universality": "they influence the everyday auditory experience of every hearing human" and other animals (Parncutt & Hair, 2011; see also Rehding, 2019). Further, harmonicity promotes consonance while roughness promotes dissonance in individual harmonic structures. Although they describe the cultural component as "simply familiarity with the music to which an individual has been exposed as well as speech and environmental sounds," there appears to

be more to it, e.g., exposure and learning, and possibly emotions as well (Parncutt & Hair, 2011; see also Arthurs, Beeston, & Timmers, 2018; Lahdelma & Eerola, 2016). However, it is not clear to what extent the two natural components versus the cultural ones affect the consonance/dissonance judgments. Furthermore, we do not have clear understanding of the roles of roughness versus harmonicity, nor of the roles of cultural components (e.g., those of familiarity or expertise versus preference) on the judgment of consonance and dissonance.

Here, *consonance* refers to a combination of tones (an interval or chord) that are perceived as concordant or in mutual agreement (*con sonare,* see, e.g., Parncutt, 1989, p. 56). *Dissonance*, in turn, refers to a combination of tones that is perceived as discordant, in mutual disagreement. We are carefully not defining dissonance *simply* as roughness (and consonance as the absence of "disturbance" or smoothness; cf. Helmholtz, 1877/1954, p. 194) or consonance as tonal fusion (and dissonance as absence of "melting"; cf. Schneider, 1997; Stumpf, 1890), nor either consonance or dissonance in terms prescribed in the practical traditions of music theory (e.g., Aldwell, Schachter, & Cadwallader, 2011, pp. 26–29; Cassiodorus, 550–562/1965, p. 89). We are also not solely defining these concepts in terms of simplicity or complexity of the (approximate) integer ratios of the fundamental frequencies of the tones (e.g., Galilei, 1638/1914, pp. 103–108; Nicomachus, as cited in Weiss & Taruskin, 1984; Parncutt & Hair, 2018). Finally, we avoid defining either term *only* based on subjective assessments such as pleasantness, tension, or other preferences, albeit some kind of judgment in consonance/dissonance is unavoidable. So, we wish to maintain conceptual possibility for the mutual interaction of two natural components (roughness and fusion) and the subjective and cultural components (such as familiarity, expertise, or preference), which all affect the perception and judgment of consonance and dissonance. We shall use the terms *sensory consonance* and *sensory dissonance* to refer to the natural components, and, when necessary, *musical consonance* and *musical dissonance* to refer to consonance/dissonance in the musical context, which is strongly linked to music cultures and affected by the individual's musical expertise and musical preferences (cf. McDermott et al., 2016; Terhardt, 1984; but see Bidelman, 2013; Bowling et al., 2017; Harrison & Pearce, 2020; Kuusi, 2001; Lahdelma & Eerola, 2016; Parncutt & Hair, 2011; Virtala & Tervaniemi, 2017). Musical consonance/dissonance is inclusive of sensory consonance/dissonance, and we acknowledge that no strict demarcation might be necessary (or even

possible) between sensory and musical consonance/dissonance, nor between natural and cultural components of perception and judgment of consonance/dissonance. Furthermore, our empirical focus is on consonance/dissonance of individual (isolated, vertical) harmonic structures, not on consonance/dissonance of their (horizontal) successions (e.g., chord progressions; cf. Lahdelma & Eerola, 2016, pp. 1–2).

The natural components of consonance/dissonance perception and judgment appear to be relatively universal (see Parncutt & Hair, 2011, p. 159). Newborn babies (Perani et al., 2010; Virtala et al., 2013) and even some non-human species (Izumi, 2000; Watanabe, Uozumi, & Tanaka, 2005) are able to discriminate consonance from dissonance. However, the cultural and subjective components rely on individual, contextual, historical, and cultural variability. Exposure to musical practices and the resulting musical enculturation leads to facilitated processing of music of one's own culture, even without explicit training (Hannon & Trainor, 2007; Virtala & Tervaniemi, 2017). The process of enculturation is modified, e.g., by motivation (Virtala & Tervaniemi, 2017), and the perception and judgment of consonance/dissonance is connected to familiarity, emotions, preferences, learning, musical expertise, and knowledge of tonal hierarchy (Arthurs et al., 2018; Johnson-Laird, Kang, & Leong, 2012; Kuusi, 2015; Lahdelma & Eerola, 2016; Omigie, Dellacherie, & Samson, 2017; Parncutt & Hair, 2011; Popescu et al., 2019; Rehding, 2019; Virtala & Tervaniemi, 2017). In particular, musicians have been shown to outperform nonmusicians in behaviorally discriminating chords and dissonance from consonance (see, e.g., Kung et al., 2014; Sares et al., 2018; Virtala et al., 2014).

BRAIN PROCESSING OF MUSICAL STIMULI

Mismatch negativity (MMN) is an event-related brain response measured with electroencephalography (EEG). A typical way to measure MMN is an oddball paradigm (Kujala, Tervaniemi, & Schröger, 2007) where a participant listens to a sound stream that consists of a frequent sound (typically 70 to 90 percent of all sounds) and an infrequent sound or sounds (typically 10 to 15 percent of all sounds, each). The infrequent sounds, or deviants, differ from the frequent sounds, or standards, in some acoustic feature like frequency, duration, or timbre. The MMN is defined as the difference in the responses between the standard and the deviant stimuli between 100–250 ms after the deviant onset (Näätänen, 1992; Näätänen, Gaillard, & Mäntysalo, 1978). The MMN originates from two sources: the bilateral supratemporal planes of the auditory cortices and the prefrontal cortex

(Näätänen & Escera, 2000; Rinne, Alho, Ilmoniemi, Virtanen, & Näätänen, 2000). The dominating theory states that the MMN is elicited when an unexpected sound occurs in a stream of expected sounds: the prediction of the oncoming sound is violated and an MMN is elicited (Näätänen, 1992; Winkler, Denham, & Nelken, 2009; for a review, see e.g., Näätänen, Paavilainen, Rinne, & Alho, 2007).

The MMN latency and amplitude are closely related to perceptual accuracy in discriminating standard and deviant stimuli: The shorter MMN latency and/or the larger the MMN amplitude, the better the perceptual accuracy for pitch (Novitski, Tervaniemi, Huotilainen, & Näätänen, 2004; Tiitinen, May, Reinikainen, & Näätänen, 1994) or duration (Amenedo & Escera, 2000). Based on this close correspondence, the MMN has been called an "objective" measure of perceptual accuracy that is elicited even when the participants are not able to attend to the stimulus (e.g., patients with deficits in consciousness or communication skills; newborn infants and toddlers). Moreover, the MMN is useful when investigating nonclinical populations as well since its presence is not affected by modulation of participants' motivation or attention. These are important issues particularly when investigating clinical populations or when comparing groups of participants who might differ in their task-specific motivation (e.g., high-level experts compared with laypeople).

Some studies have suggested that the MMN amplitude is affected by attention. Typically, the MMN is measured in a passive condition where the subject is not attending to the sound stream. However, the response has been shown to be larger in active compared to passive conditions for frequency deviations (Trejo, Ryan-Jones, & Kramer, 1995), intensity deviations in dichotic conditions (Woldorff, Hackley, & Hillyard, 1991), and violations in sound patterns (Alain & Woods, 1997). However, the MMN response in attended conditions is sometimes difficult to differentiate from the overlapping responses, such as N2b. This response occurs at the same latency but differs from the MMN e.g., based on scalp topography or sensitivity to attentional manipulation (Novak, Ritter, & Vaughan, 1992; Näätänen, Simpson, & Loveless, 1982), and in some studies where MMN and N2b have been carefully dissociated, there seems to be no sign of an effect of attention on the MMN amplitude (Gomes et al., 2000; Sussman et al., 2004; Sussman, Winkler, Huotilainen, Ritter, & Näätänen, 2002). According to Sussman (2007, 2017) the primary factor that modulates MMN is not attention to deviant stimuli, but the auditory context where the stimuli are presented; for example, some

studies have reported that if attention is focused strongly on another simultaneous sound stream, the MMN responses are attenuated (Sussman, Winkler, & Wang, 2003; Woldorff et al., 1991). Instead, Näätänen et al. (2007) suggested in their review that attentional effects on the MMN might depend on the magnitude of the stimulus change or the quality of the change: smaller changes may be more affected by the attention than large changes. Overall, the role of attention on the MMN is not yet decided unequivocally.

The MMN has been used as a tool for inspecting change detection in auditory perception in numerous studies, and its response to deviations in a sound stream appears to be a very robust phenomenon. For example, the MMN has been found to be elicited by changes in pitch (Näätänen, Pakarinen, Rinne, & Takegata, 2004; Putkinen, Tervaniemi, Saarikivi, de Vent, & Huotilainen, 2014; Tervaniemi, Huotilainen, & Brattico, 2014; Vuust et al., 2011), sound duration (Näätänen et al., 2004; Putkinen et al., 2014), intensity (Näätänen et al., 2004; Putkinen et al., 2014; Vuust et al., 2011), timbre (Christmann, Lachmann, & Berti, 2014; Putkinen et al., 2014; Vuust et al., 2011), and rhythm (Putkinen et al., 2014; Vuust et al., 2011), as well as chord structure (Virtala et al., 2014). More complex deviations, such as change in the musical rules (like the change in the direction of the sound pairs, or adding an out-of-key tone to the melody) also elicit an MMN response (see, Näätänen et al., 2007; Paavilainen, Simola, Jaramillo, Näätänen, & Winkler, 2001). Particularly, the MMN response (or its magnetic counterpart MMNm) has been detected for discriminating consonant from dissonant intervals (Crespo-Bojorque, Monte-Ordoño, & Toro, 2018; Itoh, Suwazono, & Nakada, 2010) and chords (Brattico et al., 2009; Virtala et al., 2011). Although some of these studies have compared the ERPs between musicians and nonmusicians (Brattico et al., 2009; Crespo-Bojorque et al., 2018; Itoh et al., 2010), detection of the degree of dissonance in a chord setting has not been investigated. Moreover, in all studies of consonance/dissonance to date, tonality and consonance/dissonance have been confounded since the chords were formed from traditional Western major/minor modes, and not from more consonant intervals, such as fourths and fifths. Thus there was always difference in the familiarity of the consonant chords (major and minor triads) and dissonant chords (e.g., triads consisting of a minor second and a diminished fifth such as C4–C♯4–G4.

MUSICAL EXPERTISE IN SOUND PROCESSING

Musicians typically outperform their nonmusician peers in behavioral sound change detection (Kung et al., 2014;

Sares et al., 2018; Tervaniemi, Just, Koelsch, Widmann, & Schröger, 2005; Virtala et al., 2014). Moreover, several electrophysiological studies have reported enhanced auditory discrimination skills in musicians (Brattico et al., 2009; Kung et al., 2014; Pantev et al., 1998; Vuust et al., 2005), while others have reported MMN responses for musically relevant stimulus change occurring only in musicians (Crespo-Bojorque et al., 2018; Virtala et al., 2014; see Putkinen & Tervaniemi, 2018) or even in musicians representing specific genre (Tervaniemi, Rytkönen, Schröger, Ilmoniemi, & Näätänen, 2001). Virtala and colleagues (2014) found that only musicians showed an MMN response to minor and inverted major chords in the context of major chords, while the study by Tervaniemi et al. (2001) suggested that only the musicians with long-term experience in playing from memory showed MMN responses to deviating melodic contours, as opposed to nonmusicians, and musicians relying on the musical score while performing. This benefit of musicianship in neural discrimination seems to be more pronounced for complex stimuli, such as small pitch deviations (Koelsch, Schröger, & Tervaniemi, 1999; Marques, Moreno, Castro, & Besson, 2007; Tervaniemi et al., 2005) or musically embedded pitch changes, compared to, e.g., changes in sinusoidal sounds (Brattico, Näätänen, & Tervaniemi, 2001; Fujioka, Trainor, Ross, Kakigi, & Pantev, 2004). Indeed, the amplitude of the MMN response has been shown to be associated with differences in performance in behavioral paradigms such as, e.g., detecting deviating sound patterns or chord structures among the frequent standard stimuli (Tervaniemi, Ilvonen, Karma, Alho, & Näätänen, 1997; Tervaniemi et al., 2001; Virtala et al., 2014; see also, Näätänen et al., 2007).

Consonant vs. dissonant chords and intervals are detected by both musicians and nonmusicians, although musicianship seems to enhance this discrimination both behaviorally (Arthurs et al., 2018; Kung et al., 2014) and neurally (Brattico et al., 2009; Crespo-Bojorque et al., 2018; Kung et al., 2014). Recently, Arthurs et al. (2018) showed that when listening to twelve different chord types, listeners' subjective ratings of consonance was associated with their level of experience in Western tonal music. Regarding brain measures, Crespo-Bojorque and colleagues (2018) did not find a difference in MMN amplitude between musicians and nonmusicians in neural discrimination of dissonant intervals in a consonant context, but only musicians showed an MMN response in a more atypical setting, namely, when consonant intervals were played in a dissonant context.

Previous studies have investigated the detection of dissonance vs. consonance with intervals (e.g., Arthurs

et al., 2018; Crespo-Bojorque et al., 2018; Itoh et al., 2010; Kung et al., 2014) or with chords within a major/minor context (e.g., Brattico et al., 2009; Virtala et al., 2011). Furthermore, several ERP studies have deployed consonant and dissonant intervals and chords in a dichotomic way, not considering this musical feature as a continuum from very dissonant to highly consonant sound combinations (e.g., Brattico et al., 2009; Kung et al., 2014; Virtala et al., 2011). We wanted to investigate the ability to differentiate varying degrees of dissonance when compared to consonant chords. In order to do this, we employed two different degrees of dissonance in the experimental setting. We used four-tone chords comprised of either highly consonant intervals or highly dissonant intervals which do not belong to major or minor modes. The two different degrees of dissonance were created by adding more or less dissonant intervals to the chords. To enhance the ecological validity of the study, we used piano tones, as opposed to sine tones, as stimuli.

We hypothesized that: 1) the musicians would show larger MMNs and more accurate behavioral detection of dissonance vs. consonance than nonmusicians; 2) increasing the degree of dissonance would be seen in MMN responses and behavioral discrimination of chords; and 3) attending to the sounds would enhance neural discrimination of consonance vs. dissonance compared to the passive condition.

Method

PARTICIPANTS

Fifteen nonmusicians (age 20–37 years, mean = 25.9, $SD = 5.2$, median = 24.0; 5 male) and sixteen musicians (age 20–32 years, mean = 25.7, $SD = 3.6$, median = 25; 6 male) participated in the study. All musicians had either graduated from (2) or were studying at university level music program (14). They had received formal instrument teaching for at least 10 years (mean = 14.2, $SD = 2.5$, median = 14.5, min = 10 max = 18), and were able to play 2 to 6 different instruments. Only classical musicians were included in the music group since the rules and conventions concerning dissonance are somewhat more distinct in classical music compared to, e.g., jazz music. Most of the nonmusicians had not received any instrument lessons (10). Five nonmusicians reported having participated in instrument lessons (1, 4, 5, 5 and 6 years) during childhood or adolescence. In order to avoid too much overlap in musical expertise in the groups, one participant reporting nine years of attendance in instrument lessons was excluded from the 16 individuals originally in the nonmusician group.

None of the nonmusicians regularly played an instrument during the experiments.

The participants were Finnish speaking, right-handed, and reported having normal hearing and normal or corrected-to-normal vision. The education level between the groups did not differ. One participant from the nonmusician group was additionally excluded from the analyses for the behavioral test due to not completing the task.

The participants signed the informed consent and were told that they had the right to stop the experiment whenever they wanted. For participation, they were given vouchers (5€/30 minutes) to be used for cultural interests. The experiments were approved by the University of Helsinki Ethical Review Board in the Humanities and Social and Behavioural Sciences, in Helsinki, Finland, and were carried out in accordance with the committee's guidelines and regulations, as well as with those of the Helsinki Declaration.

STIMULI

The chords of the experiment were created with Steinway piano sounds from the University of McGill DVD sound library (Opolko & Wapnick, 2006). Standard stimuli were consonant (CON) chords (dyads with octave doubles), and two types of dissonant chords (triads with an octave double of the root and tetrads) acted as deviant stimuli (D1 and D2). The chords were constructed with four tones so that: 1) they would pare down the effects of cultural context and conventions by excluding major or minor thirds that are typically included in chords of Western tonal music; 2) the ambitus, that is, the span between the outer voices (top and bottom tones of chords) does not change as the chord type changes, in order to minimize any perceptual issues relating to melodic motions and voice leading; 3) the changes between the three chord structures are produced by only a half-step change in one tone, again to minimize larger melodic movement; 4) the consonant chord consists only of the most consonant intervals (octave, perfect fifth, and perfect fourth; see e.g. Bowling & Purves, 2015); 5) the dissonant chords contain the clearly dissonant intervals of tritones and minor ninths (= minor second's octave double), *without* emphasis on roughness as a factor contributing to dissonance (avoiding intervals of minor and major second); and 6) the other dissonant chord contains one dissonant interval more (a minor ninth) and one consonant interval less (an octave) than the other dissonant chord.

Thus, and taking octave equivalents into account, the consonant chord (CON) was comprised of tones C3-G3-C4-G4, which contains 2 octaves (C3-C4 and

G3-G4), 3 perfect fifths (C3-G3, C4-G4 and C3-G4), and 1 perfect fourth (G3-C4). The first dissonant chord (D1) consisted of C3-F♯3-C4-G4, containing 1 perfect octave (C3-C4) and 2 perfect fifths (C4-G4 and C3-G4) but also 2 tritones (C3-F♯3 and F♯3-C4) and 1 minor ninth (F♯3-G4). The second dissonant chord (D2) consisted of C3-F♯3-C♯4-G4, containing no perfect octaves but 2 perfect fifths (C4-G4 and F♯3-C♯4) as consonant intervals, 2 tritones (C3-F♯3 and F♯3-C4) and 2 minor ninths (C3-C♯4 and F♯3-G4) as dissonant intervals. In D1, the fundamental frequencies of tones C3-C4-G4 are the first, second and third partial of the harmonic series starting from the fundamental frequency of C3, while the F♯3 stands apart from the series. In D2, the fundamentals of C3 and G4 are again the first and third partial of the same harmonic series on C3, while the fundamentals of F♯3 and C♯4 are the second and third partial of the series starting on the fundamental of F♯2.

Equal tempered natural-like piano sounds were used for all chords. The chords were further transposed to six different pitch levels, lowest tones being C3, D3, E3, F♯3, A♭3, and B♭3, fundamental frequency of all tones varying between C3 and F5, or 130.8–698.5 Hz. All the chords are depicted and listed in the Figure 1.

The same stimuli were used in both the active and passive conditions in the oddball paradigm (Figure 2) and in the behavioral test. The intensity of the stimuli was ∼60 dB SPL as measured with a sound level meter (Extech HD600, Extech Instruments, Boston, MA). Both EEG conditions were divided into four blocks, each including 360 stimuli. In the EEG paradigm, 80% of the stimuli were standards and each of the two deviant stimuli occurred in 10% of the stimuli. The stimuli were semi-randomized, so that two deviating stimuli never occurred consecutively, and two to six standard stimuli always occurred in succession before the next deviant. In addition, the transposition changed after each chord. The duration of the chord was 650 ms and the interstimulus interval was 350 ms.

The stimuli in the active condition (Figure 2B) was the same as in passive condition (Figure 2A) with the exception that the sound stream included also 16 violin tones (C4, 261.6 Hz). The participants were instructed to push a button when hearing a violin tone. The stimulus coming after the violin tone was excluded from the analyses to exclude the artefacts caused by muscle activity.

In a separate behavioral task, the participants were asked to discriminate between the chord types: 144 chord pairs were presented with an interstimulus interval of 800 ms between the chords of one chord pair. The participants were given 3700 ms to answer which of the two chords were more dissonant, and the next chord pair was presented automatically. The time given for answering was based on piloting the needed time to answer the task without prolonging the test unnecessarily. The chord types in one pair were always different. In half of the chord pairs (72), two dissonant chords of different types (D1 vs. D2) were presented, and in half, the consonant chord (CON) was compared to either the less dissonant chord CON vs. D1) or to the more dissonant chord (CON vs. D2). The order of the chords within chord pairs was counterbalanced and the transpositions and the order of the chord pairs were randomized.

PROCEDURE

The data were collected in the laboratory of the Department of Psychology and Logopedics at the University of Helsinki. During the experiment, the participants sat in an electrically shielded room. They were instructed to sit with their eyes open and avoid unnecessary movement. EEG measurement started with the passive condition where the participants were instructed to ignore the sounds coming from the headphones and concentrate on a muted self-chosen movie with subtitles. In the subsequent active condition, the participants looked at the picture on the screen and were instructed to concentrate on the sounds and push the button whenever they heard the violin tone. The stimuli were presented via Sony Professional MDR-7506 headphones. Each of the eight blocks lasted for 6 minutes and the participants were given the opportunity to have a break between the blocks.

The behavioral task was presented after the active EEG condition. The participants were presented with two consecutive chords and were asked to decide which of the chords was more dissonant and push the button accordingly (left-hand button for the first chord, right-hand button for the second chord). The instructions showed on the screen during the entire test. Before the test, the participants rehearsed the task: the dissonance and consonance of the chords was explained to the nonmusicians by using Finnish words for smoothness/accordance and roughness/discordance. All the participants were then played two examples of all three chord types. After this they were further played three chord pairs (consonant-more dissonant, less dissonant-more dissonant and more dissonant-less dissonant) with the experimenter giving the correct answer after each pair of which of the two chords was more dissonant. The participants were then allowed to listen to the chord pairs as many times as they wished to make sure that they understood the concepts of dissonance and consonance. The experiment lasted for 14 minutes and it was

A

B

Transposition	Consonant chord	Less dissonant chord	More dissonant chord
C	C3 - G3 - C4 - G4	C3 - F♯3 - C4 - G4	C3 - F♯3 - C♯4 - G4
D	D3 - A3 - D4 - A4	D3 - A♭3 - D4 - A4	D3 - A♭3 - E♭4 - A4
E	E3 - H3 - E4 - H4	E3 - B♭3 - E4 - H4	E3 - B♭3 - F4 - H4
F♯	F♯3 - C♯4 - F♯4 - C♯5	F♯3 - C4 - F♯4 - C♯5	F♯3 - C4 - G4 - C♯5
A♭	A♭3 - E♭4 - A♭4 - E♭5	A♭3 - D4 - A♭4 - E♭5	A♭3 - D4 - A4 - E♭5
B♭	B♭3 - F4 - B♭4 - F5	B♭3 - E4 - B♭4 - F5	B♭3 - E4 - H4 - F5

FIGURE 1. The chords used in the experiments (CON = consonant chord, D1 = less dissonant chord, D2 = more dissonant chord).

divided into two blocks, allowing the subject to have a break in between the blocks. The whole measurement (EEG and behavioral task) took 2.5 to 3 hours, total.

DATA RECORDING AND PROCESSING

The EEG was recorded with 64 Ag-AgCl scalp electrodes according to the international 10–20 system (Electro-Cap International Inc., Eaton, OH, United States) and the data were registered with Biosemi 7.07 (BioSemi B. V., Amsterdam, Netherlands) with the sample rate of 512 Hz. In addition to 64 scalp electrodes, four additional electrodes were placed: one on the nose

as a reference electrode, one below the left eye and one behind each mastoid bone. The stimuli were presented with Presentation 17.2 software (Neurobehavioral Systems, Inc., Albany, CA) and the responses in the active EEG paradigm and the behavioral test were recorded with the Cedars Response Pad (Model RB-834; Cedrus Corporation, San Pedro, CA, United States).

EEG data were processed with BESA 6.0. software (MEGIS Software GmbH, Gräfelfing, Germany). Noisy electrodes showing atypical signals (detected by visual inspection) were interpolated; that is, their signal was replaced by the average signal calculated (by BESA 6.0
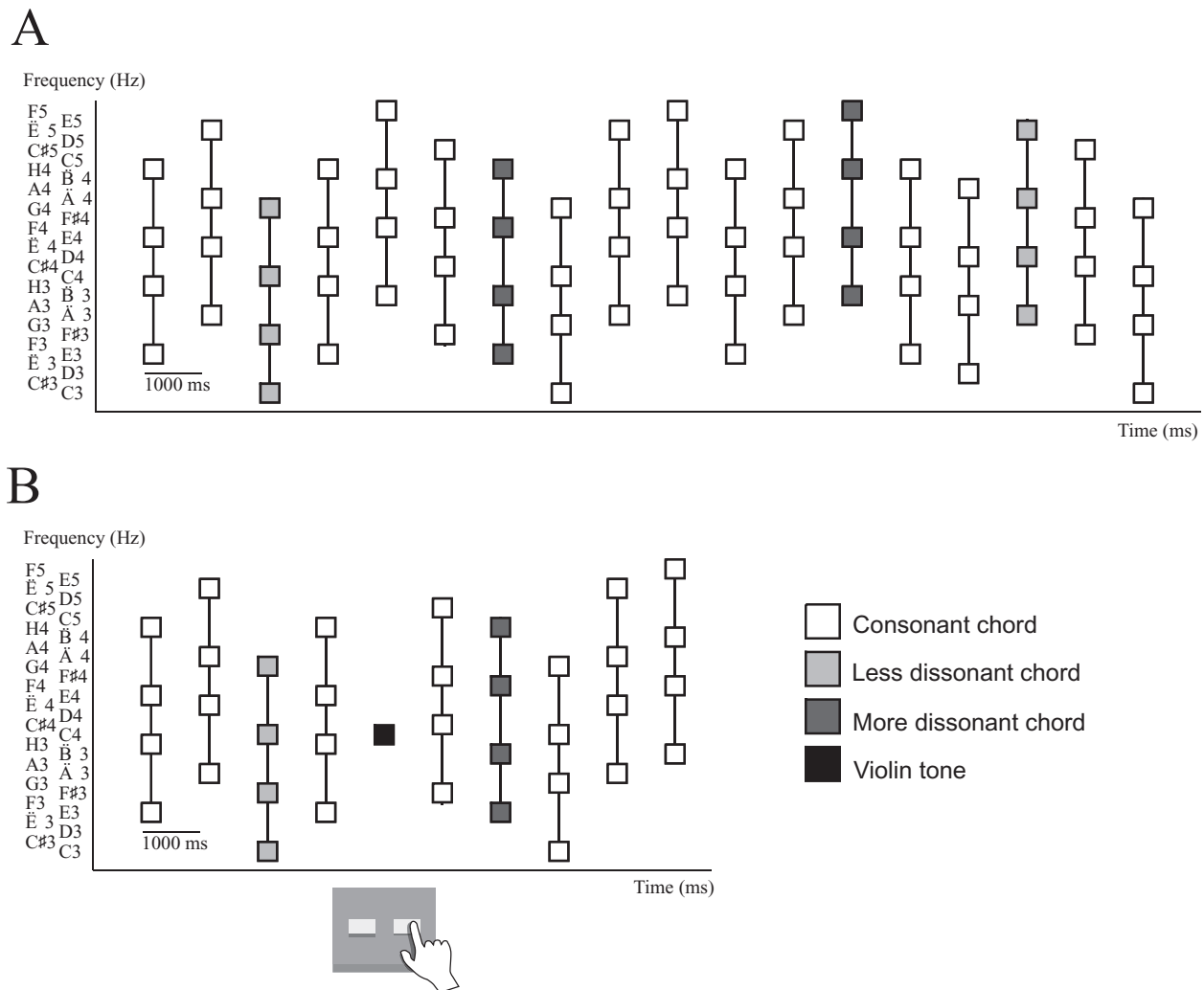
FIGURE 2. Display for the EEG paradigm. A) Passive oddball paradigm during which the participants were watching a muted self-chosen video, B) active oddball paradigm during which the participants were instructed to push a button when they heard a violin sound. Figure layout based on Virtala et al. (2014).

software) over the surrounding electrodes. We further removed eye blink artefacts exceeding ±100 μV with semi-automatic Besa PCA method. EEG-epochs with amplitudes exceeding ± 100 μV were excluded from the analyses, and the percentage of accepted trials averaged over all participants and blocks was 92.2% ($SD = 5.3$). Of the six electrodes that were used in the conclusive analyses, on average 0.11 ($SD = 0.32$, median = 0, min = 0, max = 1) were interpolated.

Frequencies below 0.5 Hz and above 30 Hz were filtered out offline and the inspected epochs were extracted from the EEG between -100 ms to 500 ms around the onset of the stimulus. The electrode signals were re-referenced to the average of the mastoid electrodes. For each participant, the responses were averaged separately over each chord type, combining all transpositions, for both active and passive conditions separately. Averaged responses were then exported to MATLAB R2016 (The MathWorks Inc., Natick, MA) for further inspection.

The subtraction signals were created for both dissonant chord types by subtracting the averaged response for the consonant chord (CON) from the averaged responses for the dissonant chords (D1 & D2) for each participant, electrode, and condition. These subtraction signals were then averaged for both groups (musicians vs. nonmusicians), and for both conditions (active vs. passive).

STATISTICAL ANALYSES

Six electrodes (F3, Fz, F4, C3, Cz, and C4) were included in the EEG analyses. The MMN responses were first identified within the typical latencies of 100–250 ms (Näätänen, 1992; Näätänen et al., 1978). Following prior literature, the mean amplitudes for MMN were then calculated over 30 ms time windows (e.g., Ilvonen et al., 2004; Leung, Croft, Baldeweg, & Nathan, 2007; Tervaniemi, Maury, & Näätänen 1994; Weise, Grimm, Müller, & Schröger, 2010), covering ±15 ms area around the peak amplitudes of the averaged subtraction waveforms on the electrode Cz, where the amplitudes were the largest. Time windows were assessed separately for each group, condition, and deviant type. We then calculated the mean MMN amplitude over 6 electrodes (F3, Fz, F4, C3, Cz, and C4) for each participant and used this mean in the analysis. These frontal and central electrodes were chosen for averaging as the MMN typically appears on these scalp regions in EEG (See Näätänen et al., 2007, for a review). We conducted a two-way [condition (active, passive) x chord (less dissonant, more dissonant)] repeated measures ANOVA (rANOVA) with group as a between-subjects factor. Greenhouse-Geisser corrections were applied in the analyses if the sphericity could not be assumed. Statistically significant main effects and interactions with more than two levels were further inspected with pairwise comparisons. To counteract the problem of multiple comparisons, Bonferroni corrections were applied.

In the behavioral test, the percentage of correct answers in three different comparison conditions (CON vs. D1; CON vs. D2; D1 vs. D2) was compared between groups with rANOVA. The percentage of correct answers was not distributed normally for all groups and comparisons: in particular, the distribution of musicians' group was skewed to the left when comparing the consonant chord to the dissonant ones. This was most likely due to a ceiling effect caused by the task being undemanding for the musicians. Because of this, the results were further inspected with the nonparametric Mann-Whitney U test. The alpha level was set at $p \leq .50$. All the analyses were conducted with IBM SPSS Statistics 25.0 (IBM Corporation, NY).

Results

EVENT-RELATED POTENTIALS

Figures 3 and 4 depict standard and deviant responses averaged over both groups in the active and passive conditions on electrodes F3, Fz, F4, C3, Cz, and C4. The subtraction signals for both deviant chords are depicted in Figures 5 and 6.

Table 1 lists the mean amplitudes and the peak latencies for dissonant chords in both conditions and both groups for MMN response for the mean amplitude over frontal and central line (C3, Cz, C4, F3, Fz, and F4). All the inspected MMN amplitudes were statistically significant, indicating that both groups detected both degrees of dissonance in passive and in active conditions.

No main effects of *group*, $F(1, 29) = 0.04$, $p = .839$, or *condition*, $F(1, 29) = 0.59$, $p = .448$, were found for MMN responses. A marginally significant main effect was found for *dissonant chord type*, $F(1, 29) = 3.67$, $p = .065$, $\eta^2 = .112$, suggesting larger mean amplitude for more—compared to less—dissonant chord type. Interaction between *group* and *condition* was significant, $F(1, 29) = 7.25$, $p = .012$, $\eta^2 = .220$, with post hoc tests indicating that nonmusicians had larger MMN responses in the passive (mean = -1.52, SEM = 0.21) than in the active (mean = -1.01, SEM = 0.23) condition ($p = .023$). The results for all main effects and interactions are listed in Table 2.

BEHAVIORAL TEST

The musicians identified the more dissonant chord accurately in 99.1% (SD = 1.3) of the CON vs. D2 chord pairs, in 96.5% (SD = 5.1) of the CON vs. D1 chord pairs and in 74.0% (SD = 8.8) of the D1 vs. D2 chord pairs. The nonmusicians were less accurate, identifying the more dissonant chord accurately in 82.1% (SD = 17.8) of the CON vs. D2 chord pairs, in 74.2% (SD = 18.2) of the CON vs. D1 chord pairs and in 63.3% (SD = 11.6) of the D1 vs. D2 chord pairs.

Group had a significant main effect on correct answers, $F(1, 28) = 19.07$, $p < .001$, $\eta^2 = .405$, with musicians giving more correct answers than nonmusicians (Figure 7A). Furthermore, there was a significant main effect of *chord pair*, $F(2, 56) = 84.06$, $p < .001$, $\eta^2 = .750$. Pairwise comparisons indicated that the fewest correct answers were given when comparing two different types of dissonant chords, followed by when comparing the less dissonant chord with the consonant chord, and the most correct answers were given when comparing the more dissonant chord with the consonant chord (D1/D2 < D1/CON, D1/CON < D2/CON; $p < .001$, each). Furthermore, there was a significant interaction of *chord pair* × *group*, $F(1.51, 54) = 5.45$, $p = .013$, $\eta^2 = .163$. Pairwise comparisons showed that the musician group discriminated both dissonant chords from the consonant chord better than they discriminated the more and less dissonant chords from each other ($p < .001$, each). In addition, they discriminated the less and the more dissonant chords from the consonant chord equally well (D1 vs. CON and D2 vs.
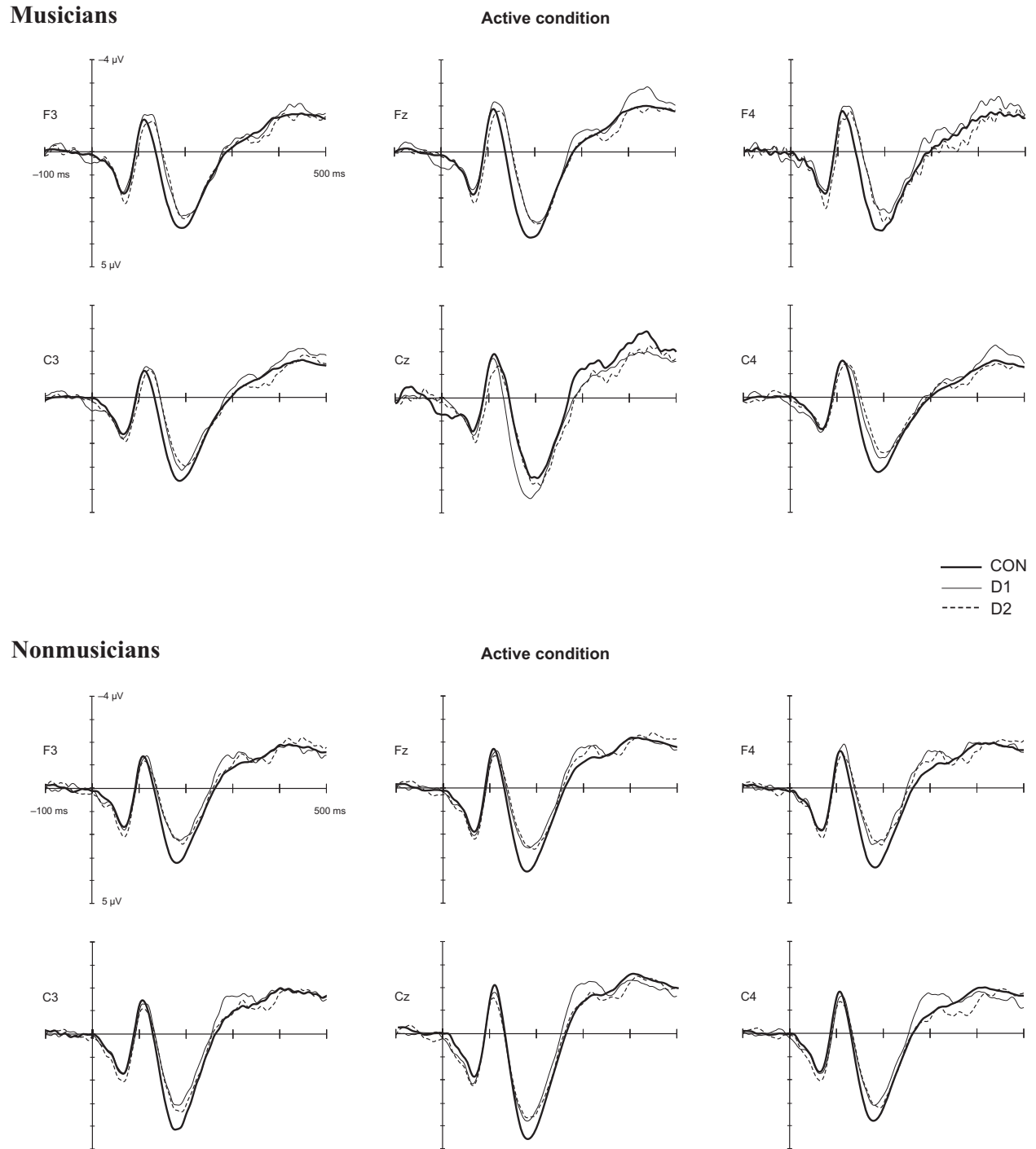
Musicians

Active condition

Nonmusicians

Active condition

CON
D1
D2

FIGURE 3. Averaged responses to consonant (CON), less dissonant (D1), and more dissonant (D2) chords for musicians (top) and nonmusicians (bottom) for the electrodes used in the analyses (F3, Fz, F4, C3, Cz, and C4) in active condition. The *x*-axes display the time from the sound onset and the *y*-axes the mean amplitude of the responses. The number of responses averaged over consonant chord is approximately eight times larger than for each dissonant chord type. Therefore, the signal-to-noise ratio differs between the responses causing a difference in their smoothness.
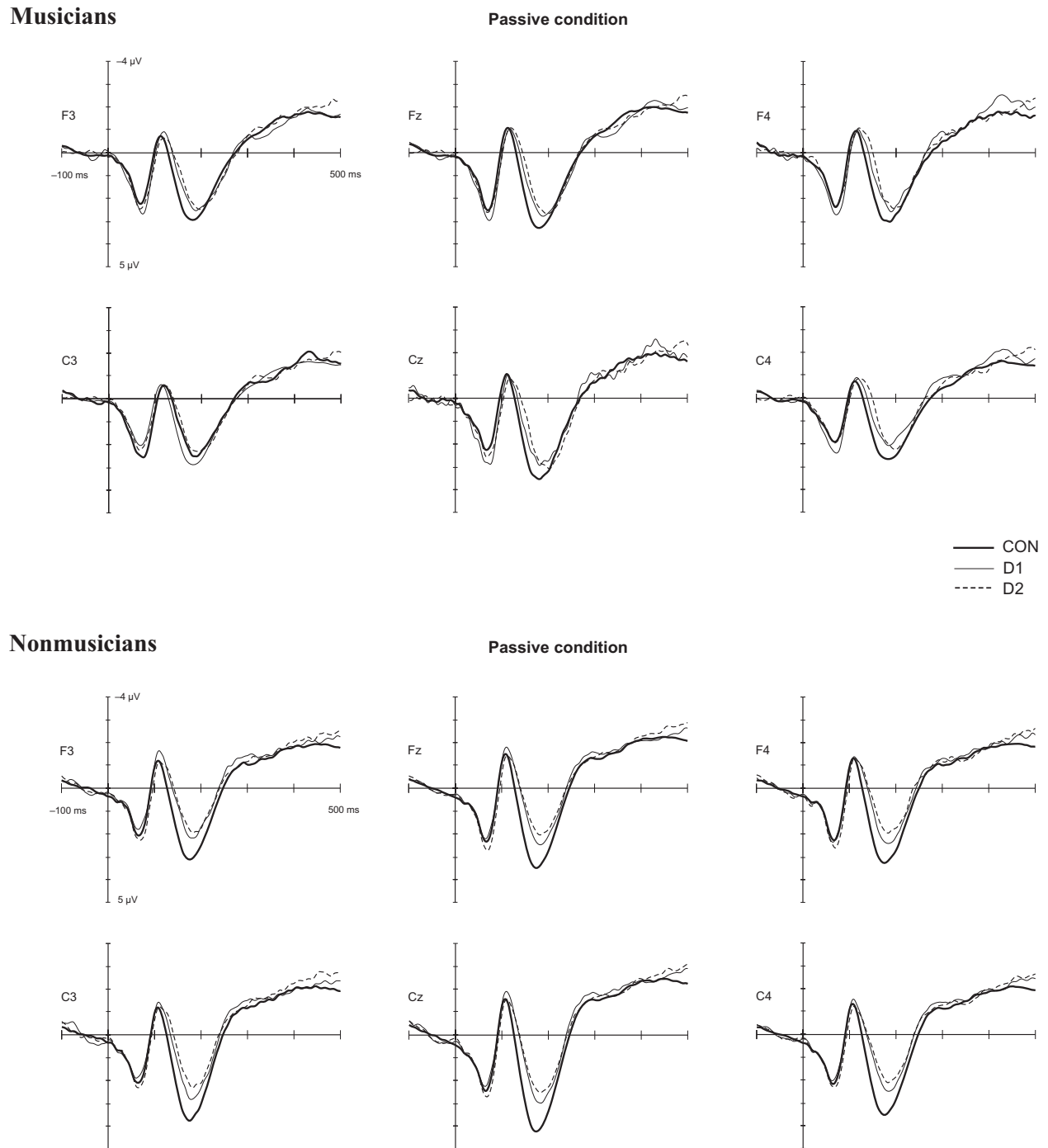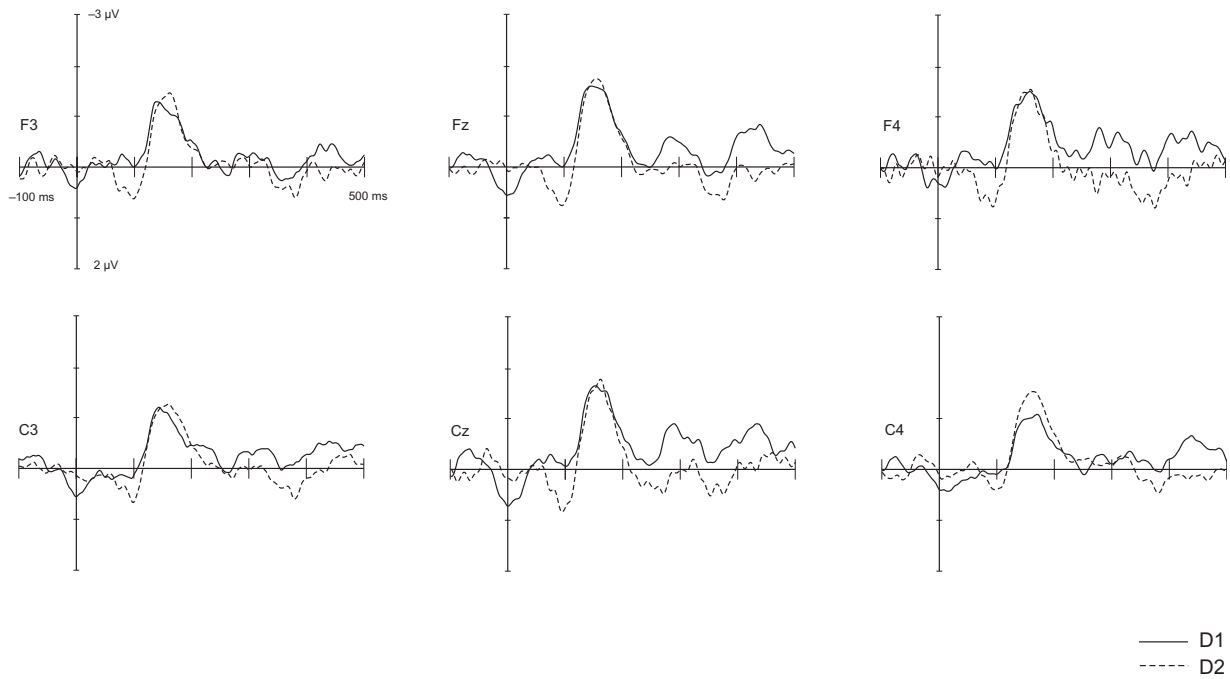
Musicians

Passive condition

Nonmusicians

Passive condition

FIGURE 4. Averaged responses to consonant (CON), less dissonant (D1), and more dissonant (D2) chords for musicians (top) and nonmusicians (bottom) for the electrodes used in the analyses (F3, Fz, F4, C3, Cz, and C4) in passive condition. The *x*-axes display the time from the sound onset and the *y*-axes the mean amplitude of the responses. The number of responses averaged over consonant chord is approximately eight times larger than for each dissonant chord type. Therefore, the signal-to-noise ratio differs between the responses causing a difference in their smoothness.

Musicians

Active condition

Nonmusicians
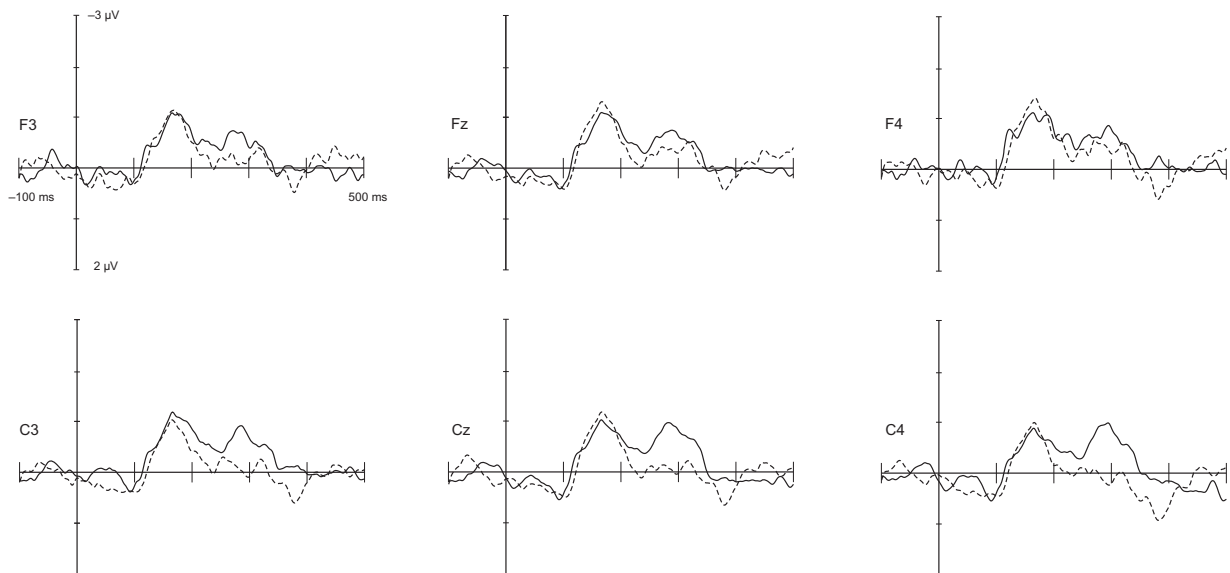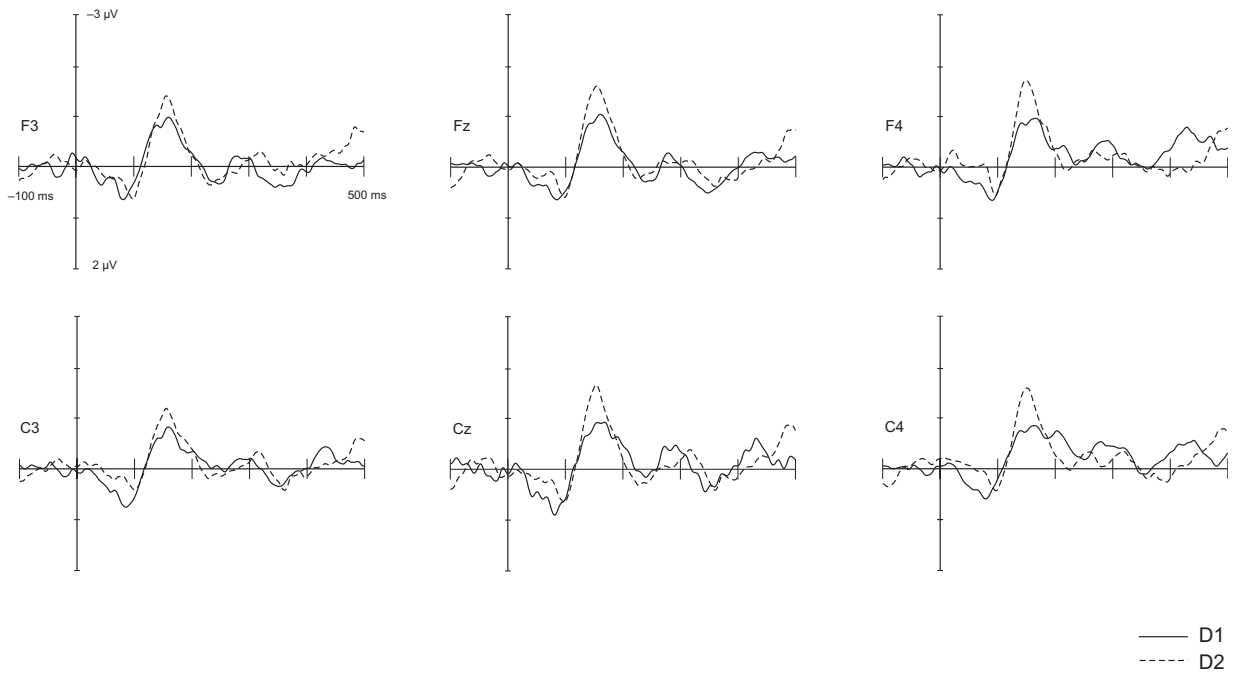
Active condition

—— D1
- - - D2

FIGURE 5. Averaged subtraction waveforms for both less dissonant (D1) and more dissonant (D2) chords for musicians (top) and nonmusicians (bottom) for the electrodes used in the analyses (F3, Fz, F4, C3, Cz, and C4) for *active conditions*. The x-axes display the time from the sound onset and the y-axes the mean amplitude of the responses. The time windows used in the analyses were assessed from the Cz, separately for both groups, and both chords: musicians: D1: 137–167 ms and D2: 147–177 ms; nonmusicians: D1:153–183 ms and D2: 153–183 ms.

Musicians

Passive condition

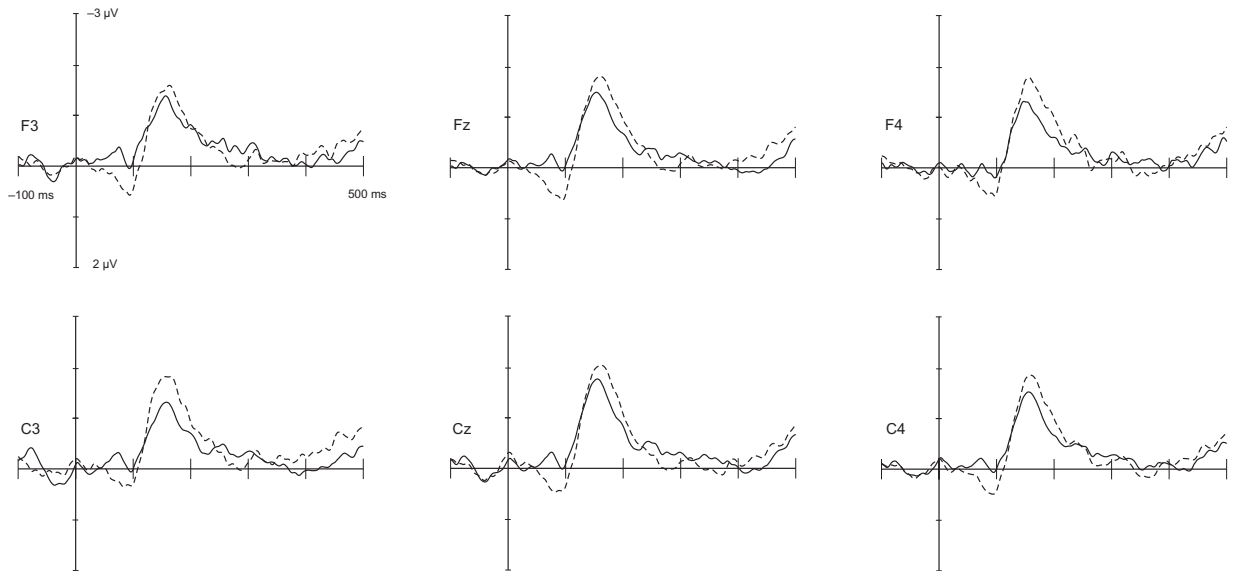**Nonmusicians**

Passive condition

FIGURE 6. Averaged subtraction waveforms for both less dissonant (D1) and more dissonant (D2) chords for musicians (top) and nonmusicians (bottom) for the electrodes used in the analyses (F3, Fz, F4, C3, Cz, and C4) for *passive conditions*. The *x*-axes display the time from the sound onset and the *y*-axes the mean amplitude of the responses. The time windows used in the analyses were assessed from the Cz, separately for both groups, and both chords: musicians: D1: 153–183 ms and D2: 139–169 ms; nonmusicians: D1:141–171 ms and D2: 149–179 ms.

TABLE 1. *Mean MMN Amplitudes Averaged Over Signals*

Condition	Musicians (*N* = 16)		Nonmusicians (*N* = 15)	
	Mean amplitude (μV) (*SD*)	Peak latency (ms)	Mean amplitude (μV) (*SD*)	Peak latency (ms)
Active				
D1	-1.29*** (0.76)	152	-0.96** (1.04)	168
D2	-1.41*** (0.91)	162	-1.05* (1.47)	168
Passive				
D1	-0.79* (1.10)	168	-1.35*** (1.04)	156
D2	-1.33*** (0.91)	154	-1.69*** (0.93)	164

Note: The mean MMN amplitudes averaged over signals from frontal and central lines (F3, Fz, F4, C3, Cz, and C4 electrode channels), along with peak latencies on the Cz electrode for less (D1) and more (D2) dissonant chords. The MMN time windows ± 15 ms around the peak amplitudes were used for the statistical analyses. The asterisks denote the statistical significance of MMN amplitude's difference from zero. *$p \leq .05$, **$p < .005$, ***$p < .001$

TABLE 2. *Results For All Main Effects and Interactions*

	F	*p*	Partial Eta Squared	Pairwise comparisons
Group	$F(1, 29) = 0.042$	$p = .839$		
Condition	$F(1, 29) = 0.593$	$p = .448$		
Dissonant chord type	$F(1, 29) = 3.669$	$p = .065$	$\eta_p^2 = .112$	
Group × Condition	$F(1, 29) = 7.249$	$p = .012$	$\eta_p^2 = .200$	Nonmusicians: Passive condition > Active condition*
Group × Dissonant chord type	$F(1, 29) = 0.159$	$p = .693$		
Condition × Dissonant chord type	$F(1, 29) = 1.370$	$p = .251$		
Group × Condition × Dissonant chord type	$F(1, 29) = 0.075$	$p = .786$		

Note: Pairwise comparisons were conducted only for the statistically significant interactions. *$p \leq .05$

CON: $p = .342$) (Figure 7B). For nonmusicians, the pairwise comparisons indicated that there were fewer correct answers when comparing two different types of dissonant chords than when comparing both more dissonant and less dissonant chords with consonant chords ($p < .001$, $p = .003$, respectively) and fewer correct answers when comparing the less dissonant chord with a consonant chord than when comparing the more dissonant chord with a consonant chord ($p < .001$). The musicians gave more correct answers in all *chord pairs* than nonmusicians (D1 vs. D2, $p = .008$; D1 vs. CON, $p < .001$; D2 vs. CON, $p = .001$).

Group comparisons were further conducted with nonparametric Mann-Whitney U test. The results were in accordance with rANOVA [CON vs. D1: U = 25.00, $p < .001$; CON vs. D2: U = 14.00, $p < .001$; D1 vs. D2: U = 51.50, $p = .012$) indicating that the groups differed from one another in all chord discrimination conditions.

Discussion

Our aim was to investigate the accuracy of neural and behavioral dissonance detection in chords with increasing degrees of dissonance. Our data indicate that both musicians and nonmusicians discriminated the dissonant chords from the consonant ones, neurally and behaviorally, but musicians outperformed nonmusicians in the behavioral task. While an increase in dissonance was detected by both groups behaviorally, the neural responses for the two types of dissonant chords differed only marginally (as reflected by the MMN). While attending to the stimuli did not modulate the musicians' MMN, the nonmusicians' MMN responses were larger in the passive than in the active condition.

BRAIN BASIS OF CONSONANCE VS. DISSONANCE DETECTION

As there was no significant difference between the MMN responses of the musician vs. nonmusician groups, our results suggest that early sensory processing of sensory consonance vs. dissonance might not be an ability requiring explicit—or maybe not even implicit—training but is based on physiological properties of auditory system. In line with this suggestion, Virtala et al. (2013) have shown that even newborn babies are sensitive to dissonant chords played among standard major chords. However, the MMN, along with the other ERP
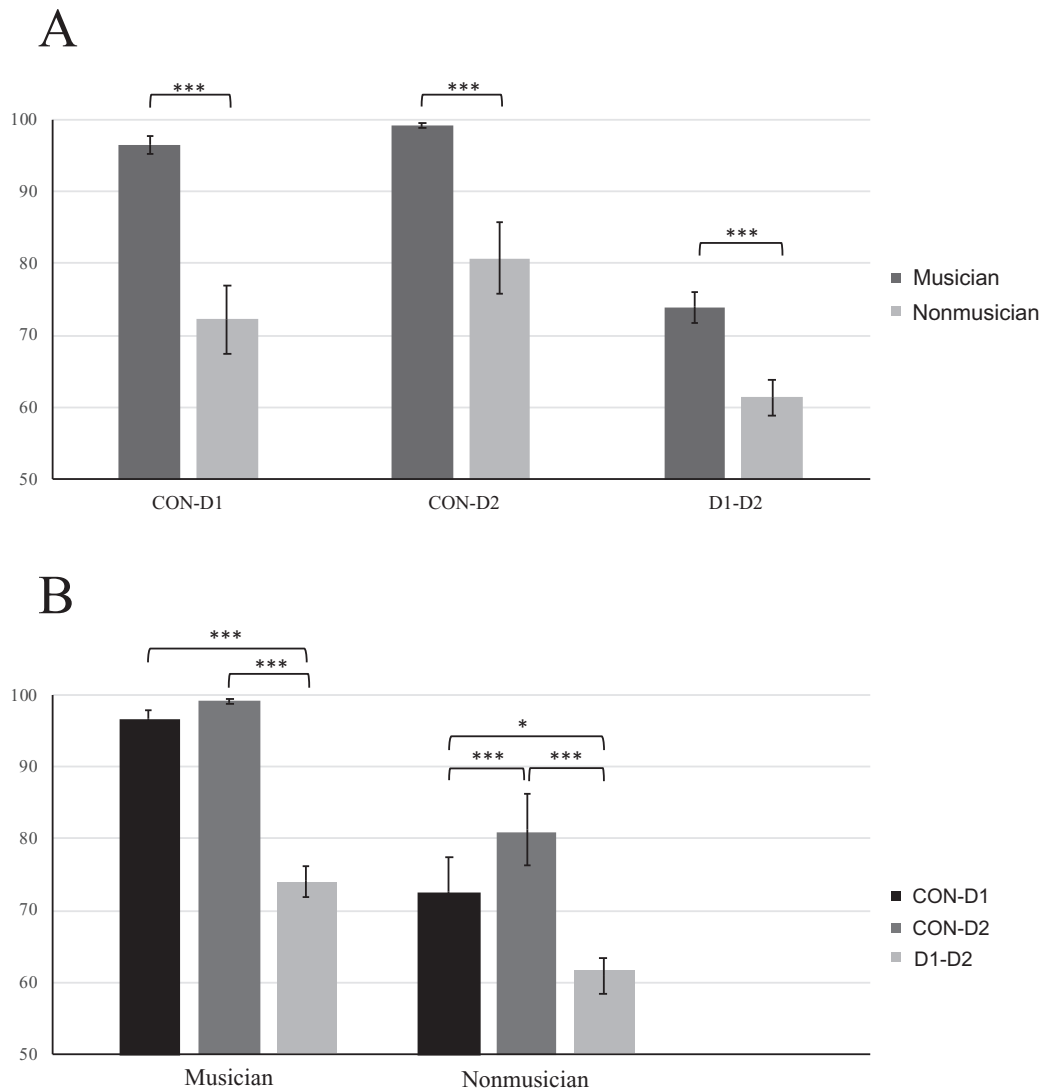
A

B

FIGURE 7. The percentage of correct answer in the behavioral discrimination task for musicians [values (SEM): CON-D1: 96.5% ($\pm$1.3%); CON-D2: 99.1% ($\pm$0.3%); D1-D2: 74.0% ($\pm$2.2%)] and nonmusicians [values (SEM): CON-D1: 72.2% ($\pm$4.8%); CON-D2: 80.7% ($\pm$4.9%); D1-D2: 61.3% ($\pm$2.6%)]. 5A) The significant differences between groups for each chord pair, 5B) the significant differences between chord pairs within both groups. CON = consonant chord, D1 = less dissonant chord, D2 = more dissonant chord; *$p \leq$ .05, **$p <$ .005, ***$p <$ .001.

components during the first 600 ms after the stimuli, is only one means to investigate neural discrimination, and thus the absence of difference between the groups' MMN responses does not necessarily indicate that early sensory detection is not modulated by expertise. Furthermore, most Western listeners are exposed to consonant and dissonant chords prior to birth, and it would be difficult—and unethical—to measure the contribution of natural components compared to that of implicit or explicit learning on dissonance detection.

There was only a marginal difference between the MMN responses for the chords. The present study

cannot determine whether this difference would become significant with a larger sample size (now $N = 31$) or not, and it does not rule out the possibility that increasing dissonance enhances its early sensory discrimination from consonance. Prior evidence shows that an increase in the degree of dissonance is detected in early auditory processing (Bidelman & Grall, 2014; Itoh et al., 2010) and further endorses the view of dissonance as a feature that is perceived as a continuum instead of dichotomic division early in the processing stream. However, based on the present study, it seems that the difference between the experts and the non-experts in the

discrimination of the degrees of dissonance is not apparent on the level of early sensory processing but only in behavioral settings where the time for sound processing is substantially longer than in typical ERP paradigms.

Previous studies have suggested that the accuracy of behavioral detection is linked to larger MMN responses (Amenedo & Escera, 2000; Jaramillo, Paavilainen, & Näätänen, 2000; Novitski et al., 2004; Tiitinen et al., 1994). Our group-level results are not in line with these studies. Further, as the behavioral test was far too easy for the musicians in the current study, it was not possible to determine whether the neural and behavioral detection of the chords correlated on an individual level. Worth noting here is that the interval structures of the dissonant chords emphasized differently not only roughness but also harmonicity: the rougher chord (D2) contained two sets of two tones that constitute consonant intervals, while the smoother chord (D1) contained one set of three tones and one tone conflicting with that harmonic series. This, along with the differences between behavioral and MMN results, might suggest that the processing of dissonance and consonance may involve factors that the MMN does not detect, such as harmonicity. The early sensory processing of degrees of dissonance calls thus for more research, with paradigms consisting of dissonant chords structured in different ways regarding roughness and harmonicity, using different chords as frequent standards and infrequent deviating stimuli, and a spectrum of event-related brain potential (ERP) methods.

It might be that different interval structures within varying chord types—instead of the difference between consonance and dissonance per se—have an impact on the MMN response. Nevertheless, previous research has shown that unlike changes in chord type, inverted major chords among root major chords do not elicit an MMN response in nonmusicians (Virtala et al., 2011; Virtala et al., 2014). In these two studies, several transpositions were used for standard and deviant chords, and all deviant chords included only tones also present in the standard chords. These results suggest that the changing interval structure within the chord does not cause a change in the MMN, at least for laypeople. As the responses did not statistically differ between the musicians and nonmusicians in our study, it seems possible that the MMN response was not evoked by the different interval structures *per se* used in the deviant chord. To further explore this, consonant chords with inversion should be added in future studies to the deviating stimuli to rule out the other interpretation.

ATTENTION AND DISSONANCE DISCRIMINATION

The present study does not support previous results suggesting that attending to stimuli elicits larger MMN amplitudes compared to the passive paradigms (Alain & Woods, 1997; Trejo et al., 1995; Woldorff et al., 1991). In contrast, the MMN response was modulated by attention in nonmusicians, who showed larger MMN amplitudes to dissonance in passive compared to active conditions. Therefore, our result might support the studies showing that focused attention attenuates the MMN response in the passively listened sound stream (Sussman et al., 2003; Woldorff et al., 1991). Although the previous studies have used dichotic listening paradigms, they do not rule out the possibility that focusing attention to the highly salient sounds within the same sound stream might affect the neural discrimination of small deviations. This is also in line with Sussman's (2007) proposal that passive processing of non-target sounds is attenuated while focusing to target sounds. In order to avoid artefact ERPs related to pressing the button in the present study, the target sound was very different from the actual stimuli, minimizing the possibility that the participants—especially the nonmusicians—would confuse target tones with the stimuli. This setting may have led the participants to ignore the actual harmonic structures of the chords and concentrate solely on the arguably more substantial timbral differences between the stimuli and target tones. Instead, due to their expertise in listening to musical sounds, the musicians might have needed to concentrate less on target sounds, so for them the difference between the conditions may not have been substantial. Finally, it cannot be ruled out that the musicians were listening covertly to the sounds during the passive condition, and this was why their brain responses did not differ between the conditions.

THE EFFECTS OF MUSICAL EXPERTISE

In line with our results, Crespo-Bojorque et al. (2018) found no differences between expert and non-expert listeners in typical Western music context. Nevertheless, they did find modulating effects of musical expertise in a less conventional context; that is, infrequent consonant intervals within standard dissonant intervals. Also, Tervaniemi et al. (2005) failed to find any difference in the MMN responses between musicians and laypeople for frequency changes (although their discrimination performance during behavioral task differed). These studies support the suggestion that neural discrimination of relatively easy musical deviation is not affected by musical expertise beyond the implicit learning through everyday exposure to Western music. It seems

thus possible that the paradigm of the present study may have been too "easy" to differentiate the groups on the early sensory processing level. The dissonant chords were not compared to each other, but both types of dissonance appeared in the middle of consonant chords.

Contradictory previous evidence using MMN as a marker for neural detection may stem from a difference in the settings: Brattico et al. (2009) used major triads and three-tone dissonant chords, whose frequency ratios do not differ from each other as much as the dissonant and consonant chords in our setting. Thus, this part of their experiment may have been more demanding for the listeners, highlighting the musician advantage. However, the results from two previous studies differ in their results for discrimination of minor and major chords. In Brattico et al. (2009), no difference was found in the groups' MMN response for the deviating minor chord. Instead, Virtala et al. (2014) found that only musicians showed MMN responses to minor and inverted major chords among major chords. This discrepancy between these two studies could be due to the subjects: Brattico et al. (2009) had smaller groups and more musical experience in the nonmusician group than in the Virtala et al. (2014) study. Moreover, only Virtala et al. (2014) used chords transposed to several frequency levels, thus making the stimulation context far more demanding than in Brattico et al. (2009).

In line with our original hypothesis, the musicians more accurately detected the dissonant chords than the nonmusicians in the behavioral task, and they also showed less within-group variability. The musicians discriminated both types of dissonant chords from the consonant one equally well. This may be explained by a ceiling effect, due to the ease of the task for the experts. The nonmusicians were less accurate. However, even the nonmusicians performed better than chance level in all discrimination tasks, showing that detecting the differences between dissonant chords is not dependent on intensive training alone.

Taken together, the group-related results in ERPs and behavioral test indicate that while expertise does not facilitate discrimination as assessed by the MMN, it does facilitate behavioral performance. Thus, the effects of musical experience appear to emerge somewhere between the latency of the MMN response ($\sim$150–250 ms) and the button press in the behavioral task ($\sim$1000–2500 ms). It may be that the better performance of the musicians in the task is also partially due to better working memory capacity in musicians (e.g., Pallesen et al., 2010; Zuk, Benjamin, Kenyon, & Gaab, 2014). This capacity could thus facilitate the conscious comparison of sounds during the behavioral task.

LIMITATIONS OF THE STUDY

The musicians in the study were all trained in Western classical music. As the individual instruments and genres may have an effect on early sensory processing, the results may not be generalizable to all musicians (Tervaniemi, Janhunen, Kruck, Putkinen, & Huotilainen, 2016; Vuust, Brattico, Seppänen, Näätänen, & Tervaniemi, 2012), or other musical cultures.

The experimental setting tested the discrimination of dissonance in a consonant context. Differences between experts and laypeople might arise in more challenging settings, such as presenting consonant or varying levels of dissonant chords in dissonant contexts (Crespo-Bojorque et al., 2018). In an ideal experiment, both contexts, and more levels of dissonant chords could be explored. Unfortunately, this was out of the scope in the current project due to lack of sufficient resources.

It cannot be ruled out that the participants were actively listening to the sounds during the passive condition, and this may explain the lack of difference between the responses in the two conditions. In particular, the musicians might have been tempted to concentrate on the sounds in the passive condition, which was always played first. This was done to avoid the learning effects affecting the passive condition. However, participants might have been more fatigued and had more difficulties in maintaining attention in the active condition.

CONCLUSIONS

The present study offers new insight about the neural basis of sensory dissonance and the effects of expertise. Furthermore, it reveals the effects of expertise on the behavioral discrimination of different degrees of dissonance. Early sensory processing of consonance vs. dissonance, as indexed by the MMN, was processed irrespective of musical expertise. Attending to the stimuli had an impact only on the nonmusician group, which showed smaller MMN amplitudes for the active than the passive condition, suggesting that focusing strongly on specific sounds attenuates processing of unessential sounds for laypeople. As the facilitating effects of expertise were apparent only in the behavioral discrimination test, it seems that the musician benefit in discrimination of consonance vs. dissonance emerges somewhere between the latency of the MMN response and the button press in the behavioral task. In addition, the study shows marginal support for the notion of dissonance as a continuum, which appears in the early stages of auditory processing.

References

ALAIN C., & WOODS D. L. (1997). Attention modulates auditory pattern memory as indexed by event-related brain potentials. *Psychophysiology, 34*, 534–546.

ALDWELL E., SCHACHTER C., & CADWALLADER A. (2011). *Harmony and voice leading* (4th ed.). Boston, MA: Schirmer.

AMENEDO E., & ESCERA C. (2000). The accuracy of sound duration representation in the human brain determines the accuracy of behavioural perception. *The European Journal of Neuroscience, 12*, 2570–2574.

ARTHURS Y., BEESTON A. V., & TIMMERS R. (2018). Perception of isolated chords: Examining frequency of occurrence, instrumental timbre, acoustic descriptors and musical training. *Psychology of Music, 46*, 662–681.

BIDELMAN G. M. (2013). The role of the auditory brainstem in processing musically relevant pitch. *Frontiers in Psychology, 4*, 264. DOI: 10.3389/fpsyg.2013.00264

BIDELMAN G., & GRALL J. (2014). Functional organization for musical consonance and tonal pitch hierarchy in human auditory cortex. *NeuroImage, 101*, 204–214.

BOWLING D. L., HOESCHELE M., GILL K. Z., & FITCH W. T. (2017). The nature and nurture of musical consonance. *Music Perception, 35*, 118–121.

BOWLING D. L., & PURVES D. (2015). A biological rationale for musical consonance. *Proceedings of the National Academy of Sciences, 112*, 11155–11160.

BRATTICO E., NÄÄTÄNEN R., & TERVANIEMI M. (2001). Context effects on pitch perception in musicians and nonmusicians: evidence from event-related-potential recordings. *Music Perception, 19*, 199–222.

BRATTICO E., PALLESEN K., VARYAGINA O., BAILEY C., ANOUROVA I., JÄRVENPÄÄ, M., et al. (2009). Neural discrimination of nonprototypical chords in music experts and laymen: An MEG study. *Journal of Cognitive Neuroscience, 21*, 2230–2244.

CASSIODORUS F. M. A. (1965). From the *Institutiones*. In O. Strunk (Ed.), *Source readings in music history 1: Antiquity and the middle ages* (pp. 87–92). London, UK: Faber and Faber. (Originally written between 550 and 562 CE)

CAZDEN N. (1980). The definition of consonance and dissonance. *International Review of the Aesthetics and Sociology of Music, 11*, 123–168.

CHRISTMANN C., LACHMANN T., & BERTI S. (2014). Earlier timbre processing of instrumental tones compared to equally complex spectrally rotated sounds as revealed by the mismatch negativity. *Neuroscience Letters, 581*, 115–119.

CRESPO-BOJORQUE P., MONTE-ORDOÑO, J., & TORO J. (2018). Early neural responses underlie advantages for consonance over dissonance. *Neuropsychologia, 117*, 188–198.

DIBBEN N. (1999). The perception of structural stability in atonal music: The influence of salience, stability, horizontal motion, pitch commonality, and dissonance. *Music Perception, 16*, 265–294.

FUJIOKA T., TRAINOR L. J., ROSS B., KAKIGI R., & PANTEV C. (2004). Musical training enhances automatic encoding of melodic contour and interval structure. *Journal of Cognitive Neuroscience, 16*, 1010–1021.

GALILEI G. (1914). *Dialogues concerning two new sciences* (H. Crew & A. de Salvio, Trans.). New York: Macmillan. (Original work published 1638)

GOMES H., MOLHOLM S., RITTER W., KURTZBERG D., COWAN N., & VAUGHAN H.G. JR. (2000). Mismatch negativity in children and adults, and effects of an attended task. *Psychophysiology, 37*, 807–816.

HANNON E. E., & TRAINOR L. J. (2007). Music acquisition: Effects of enculturation and formal training on development. *Trends in Cognitive Sciences, 11*, 466–472. DOI: 10.1016/j.tics.2007.08.008

HARRISON P. M. C., & PEARCE M. T. (2020). Simultaneous consonance in music perception and composition. *Psychological Review, 127*, 216–244. http://dx.doi.org/10.1037/rev0000169

HELMHOLTZ H. L. F. VON (1954). *On the sensations of tone* (A. J. Ellis, Trans.). New York: Dover. (Original work published 1877)

ILVONEN T., KUJALA T., KOZOU H., KIESILÄINEN A., SALONEN O., ALKU P., & NÄÄTÄNEN R. (2004). The processing of speech and non-speech sounds in aphasic patients as reflected by the mismatch negativity (MMN). *Neuroscience Letters, 366*, 235–240. https://doi.org/10.1016/j.neulet.2004.05.024

ITOH K., SUWAZONO S. & NAKADA T. (2010). Central auditory processing of noncontextual consonance in music: An evoked potential study. *Journal of the Acoustical Society of America, 128*, 3781–3787.

Izumi A. (2000). Japanese monkeys perceive sensory consonance of chords. *Journal of the Acoustical Society of America, 108,* 3073–3078.

Jaramillo M., Paavilainen P., & Näätänen R. (2000). Mismatch negativity and behavioural discrimination in humans as a function of the magnitude of change in sound duration. *Neuroscience Letters, 290,* 101–104.

Johnson-Laird P. N., Kang O. E., & Leong Y. C. (2012). On musical dissonance. *Music Perception, 30,* 19–35. DOI: 10.1525/mp.2012.30.1.19

Koelsch S., Schröger E., & Tervaniemi M. (1999). Superior pre-attentive auditory processing in musicians. *Neuroreport, 10,* 1309–1313.

Kopp D. (1995). On the function of function. *Music Theory Online, 1*(3).

Kujala T., Tervaniemi M., & Schröger E. (2007). The mismatch negativity in cognitive and clinical neuroscience: Theoretical and methodological considerations. *Biological Psychology, 74,* 1–19.

Kung C., Hsieh T., Liou J., Lin K., Shaw F., & Liang S. (2014). Musicians and non-musicians' different reliance of features in consonance perception: A behavioral and ERP study. *Clinical Neurophysiology, 125,* 971–978.

Kuusi T. (2001). *Set-class and chord: Examining connection between theoretical resemblance and perceived closeness* (Doctoral dissertation). Sibelius Academy, Helsinki, Finland. Retrieved from https://helda.helsinki.fi/handle/10138/235090

Kuusi T. (2015). Musical training and musical ability: Effects on chord discrimination. *Psychology of Music, 43,* 291–301.

Lahdelma I., & Eerola T. (2016). Mild dissonance preferred over consonance in single chord perception. *i-Perception, 7,* 1–21.

Leung S., Croft R. J., Baldeweg T., & Nathan P. J. (2007). Acute dopamine D1 and D2 receptor stimulation does not modulate mismatch negativity (MMN) in healthy human subjects. *Psychopharmacology, 194,* 443–451. https://doi.org/10.1007/s00213-007-0865-1

Marques C., Moreno S., Castro S. L., & Besson M. (2007). Musicians detect pitch violation in a foreign language better than nonmusicians: Behavioral and electrophysiological evidence. *Journal of Cognitive Neuroscience, 19,* 1453–63.

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

Novak G., Ritter W., & Vaughan H. G. Jr. (1992). The chronometry of attention-modulated processing and automatic mismatch detection. *Psychophysiology, 29,* 412–430.

Novitski N., Tervaniemi M., Huotilainen M., & Näätänen R. (2004). Frequency discrimination at different frequency levels as indexed by electrophysiological and behavioral measures. *Cognitive Brain Research, 20,* 26–36.

Näätänen R. (1992). *Attention and brain function.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Näätänen R., & Escera C. (2000). Mismatch negativity: Clinical and other applications. *Audiology Neuro Otology, 5,* 105–110.

Näätänen R., Gaillard A., & Mäntysalo S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica, 42,* 313–329.

Näätänen R., Paavilainen P., Rinne T., & Alho K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology, 118,* 2544–2590.

Näätänen R., Pakarinen S., Rinne T., & Takegata R. (2004). The mismatch negativity (MMN): Towards the optimal paradigm. *Clinical Neurophysiology, 115,* 140–144.

Näätänen R., Simpson M., & Loveless N. E. (1982). Stimulus deviance and evoked potentials. *Biological Psychology, 14,* 53–98.

Omigie D., Dellacherie D., & Samson S. (2017). Effects of learning on dissonance judgments. *Journal of Interdisciplinary Music Studies, 8,* 11–12.

Opolko F., & Wapnick J. (2006). *McGill University master samples collection on DVD.* Montreal, Quebec: McGill University.

Paavilainen P., Simola J., Jaramillo M., Näätänen R., & Winkler I. (2001). Preattentive extraction of abstract feature conjunctions from auditory stimulation as reflected by the mismatch negativity (MMN). *Psychophysiology, 38,* 359–365.

Pallesen K.J., Brattico E., Bailey C.J., Korvenoja A., Koivisto J., Gjedde A., & Calrson S. (2010). Cognitive control in auditory working memory is enhanced in musicians. *PLOS ONE, 5,* e11120.

Pantev C., Oostenveld R., Engelien A., Ross B., Roberts L., & Hoke M. (1998). Increased auditory cortical representation in musicians. *Nature, 392,* 811–814.

Parncutt R. (1989). *Harmony: A psychoacoustical approach.* Berlin: Springer-Verlag.

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

Parncutt R., & Hair G. (2018). A psychocultural theory of musical interval: Bye bye Pythagoras. *Music Perception, 35,* 475–501.

Perani D., Saccuman M.C., Scifo P., Spada D., Andreolli G., Rovelli R., et al. (2010). Functional specializations for music processing in the human newborn brain. *Proceedings of the National Academy of Sciences, 107,* 4758–4763.

Popescu T., Neuser M. P., Neuwirth M., Bravo F., Mende W., Boneh O., et al. (2019). The pleasantness of sensory dissonance is mediated by musical style and expertise. *Scientific reports, 9,* 1–11.

PUTKINEN V., & TERVANIEMI M. (2018). Neuroplasticity in music learning. In M. H. Thaut, & D. A. Hodges (Eds.), *The Oxford handbook of music and the brain* (Oxford handbooks online). Oxford, UK: Oxford University Press. doi.org/10.1093/oxfordhb/9780198804123.013.22

PUTKINEN V., TERVANIEMI M., SAARIKIVI S., DE VENT N., & HUOTILAINEN M. (2014). Investigating the effects of musical training on functional brain development with a novel Melodic MMN paradigm. *Neurobiology of Learning and Memory, 110,* 8–15.

REHDING A. (2003). *Hugo Riemann and the birth of modern musical thought.* Cambridge, UK: Cambridge University Press.

REHDING A. (2019). Consonance and dissonance. In A. Rehding & S. Rings (Eds.), *The Oxford handbook of critical concepts in music theory.* Oxford, UK: Oxford University Press. DOI: 10.1093/oxfordhb/9780190454746.013.14

RINNE T., ALHO K., ILMONIEMI R. J., VIRTANEN J., & NÄÄTÄNEN R. (2000). Separate time behaviors of the temporal and frontal mismatch negativity sources. *NeuroImage,12,* 14–19.

SARES A., FOSTER N., ALLEN K., & HYDE K. (2018). Pitch and time processing in speech and tones: The effects of musical training and attention. *Journal of Speech, Language, and Hearing Research, 61,* 496–509.

SCHNEIDER A. (1997). "Verschmelzung", tonal fusion, and consonance: Carl Stumpf revisited. In M. Leman (Ed.), *Music, gestalt, and computing.* Berlin, Heidelberg: Springer.

SUSSMAN E. (2007). A new view on the MMN and attention debate: The role of context in processing auditory events. *Journal of Psychophysiology, 21,* 164–175.

SUSSMAN E. S. (2017). Auditory scene analysis: An attention perspective. *Journal of Speech, Language, and Hearing Research, 60,* 2989–3000.

SUSSMAN E., KUJALA T., HALMETOJA J., LYYTINEN H., ALKU P., & NÄÄTÄNEN R. (2004). Automatic and controlled processing of acoustic and phonetic contrasts. *Hearing Research, 190,* 128–140.

SUSSMAN E., WINKLER I., HUOTILAINEN M., RITTER W., & NÄÄTÄNEN R. (2002). Top-down effects can modify the initially stimulus-driven auditory organization. *Cognitive Brain Research, 13,* 393–405.

SUSSMAN E., WINKLER I., & WANG W. J. (2003). MMN and attention: Competition for deviance detection. *Psychophysiology, 40,* 430–435.

STUMPF C. (1890). *Tonpsychologie.* Leipzig: S. Hirzel.

TENNEY J. (1988). *A history of 'consonance' and 'dissonance.'* New York: Excelsior.

TERHARDT E. (1984). The concept of musical consonance: A link between music and psychoacoustics. *Music Perception, 1,* 276–295.

TERVANIEMI M., HUOTILAINEN M., & BRATTICO E. (2014). Melodic multi-feature paradigm reveals auditory profiles in music-sound encoding. *Frontiers in Human Neuroscience, 8*(496). DOI: 10.3389/fnhum.2014.00496

TERVANIEMI M., ILVONEN T., KARMA K., ALHO K., & NÄÄTÄNEN R. (1997). The musical brain: Brain waves reveal the neurophysiological basis of musicality in human subjects. *Neuroscience Letters, 226,* 1–4.

TERVANIEMI M., JANHUNEN L., KRUCK S., PUTKINEN V., & HUOTILAINEN M. (2016). Auditory profiles of classical, jazz, and rock musicians: Genre-specific sensitivity to musical sound features. *Frontiers in Psychology, 6*(1900).

TERVANIEMI M., JUST V., KOELSCH S., WIDMANN A., & SCHRÖGER E. (2005). Pitch discrimination accuracy in musicians vs nonmusicians: An event-related potential and behavioral study. *Experimental Brain Research, 161,* 1–10.

TERVANIEMI M., MAURY S., & NÄÄTÄNEN R. (1994). Neural representations of abstract stimulus features in the human brain as reflected by the mismatch negativity. *Neuroreport, 5,* 844–846.

TERVANIEMI M., RYTKÖNEN M., SCHRÖGER E., ILMONIEMI R. J., & NÄÄTÄNEN R. (2001). Superior formation of cortical memory traces for melodic patterns in musicians. *Learning and Memory, 8,* 295–300.

TIITINEN H., MAY P., REINIKAINEN K., & NÄÄTÄNEN R. (1994). Attentive novelty detection in humans is governed by preattentive sensory memory. *Nature, 372,* 90–92.

TREJO L. J., RYAN-JONES D. L., & KRAMER A. F. (1995). Attentional modulation of the mismatch negativity elicited by frequency differences between binaurally presented tone bursts. *Psychophysiology, 32,* 319–328.

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,* 406–410.

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.

VIRTALA P., HUOTILAINEN M., PARTANEN E., & TERVANIEMI M. (2014). Musicianship facilitates the processing of western music chords - An ERP and behavioral study. *Neuropsychologia, 61,* 247–258.

VIRTALA P., HUOTILAINEN M., PUTKINEN V., MAKKONEN T., & TERVANIEMI M. (2012). Musical training facilitates the neural discrimination of major versus minor chords in 13-year-old children. *Psychophysiology, 49,* 1125–1132.

VIRTALA P., & TERVANIEMI M. (2017). Neurocognition of major-minor and consonance-dissonance. *Music Perception, 34,* 387–404.

VUUST P., BRATTICO E., GLEREAN E., SEPPÄNEN M., PAKARINEN S., TERVANIEMI M., & NÄÄTÄNEN R. (2011). New fast mismatch negativity paradigm for determining the neural prerequisites for musical ability. *Cortex, 47,* 1091–1098.

Vuust P., Brattico E., Seppänen M., Näätänen R., & Tervaniemi M. (2012). The sound of music: Differentiating musicians using a fast, musical multi-feature mismatch negativity paradigm. *Neuropsychologia, 50*, 1432–1443.

Vuust P., Pallesen K. J., Bailey C., van Zuijen T. L., Gjedde A., Roepstorff A., & Østergaard L. (2005). To musicians, the message is in the meter: Pre-attentive neuronal responses to incongruent rhythm are left-lateralized in musicians. *NeuroImage, 24*, 560–564.

Watanabe S., Uozumi M., & Tanaka N. (2005). Discrimination of consonance and dissonance in Java sparrows. *Behavioural Processes, 70*, 203–208.

Weise A., Grimm S., Müller D., & Schröger E. (2010). A temporal constraint for automatic deviance detection and object formation: A mismatch negativity study. *Brain Research, 1331*, 88–95.

Weiss, P., & Taruskin, R. (1984). *Music and the Western world.* New York: Schirmer Books.

Winkler I., Denham S. L., & Nelken I. (2009). Modeling the auditory scene: Predictive regularity representations and perceptual objects. *Trends in Cognitive Sciences, 13*, 532–540.

Woldorff M. G., Hackley S. A., & Hillyard S. A. (1991). The effects of channel-selective attention on the mismatch negativity wave elicited by deviant tones. *Psychophysiology, 28*, 30-42.

Zarlino G. (1558) [1982]. From the Istitutione armoniche. In O. Strunk (Ed.), *Source readings in music history 2: Renaissance* (pp. 38–71). London: Faber and Faber.

Zuk J., Benjamin C., Kenyon A., & Gaab N. (2014). Behavioral and neural correlates of executive functioning in musicians and non-musicians. *PLOS One* 9, e99868.