A Psychocultural Theory of Musical Interval: Bye Bye Pythagoras

Richard Parncutt
University of Graz, Graz, Austria

Graham Hair
Manchester Metropolitan University, Manchester, United Kingdom & University of Glasgow, Glasgow, United Kingdom

The Pythagoreans linked musical intervals with integer ratios, cosmic order, and the human soul. The empirical approach of Aristoxenus, based on real musicians making real music, was neglected. Today, many music scholars and researchers still conceptualize intervals as ratios. We argue that this idea is fundamentally incorrect and present convergent evidence against it. There is no internally consistent "Just" scale: a 6th scale degree that is 5:3 above the 1st is not a perfect 5th (3:2) above the 2nd (9:8). Pythagorean tuning solves this problem, but creates another: ratios of psychologically implausible large numbers. Performers do not switch between two ratios of one interval (e.g., 5:4 and 81:64 for the major third), modern studies of performance intonation show no consistent preferences for specific ratios, and no known brain mechanism is sensitive to ratios in musical contexts. Moreover, physical frequency and perceived pitch are not the same. Rameau and Helmholtz derived musical intervals from the harmonic series, which is audible in everyday sounds including voiced speech; but those intervals, like musical intervals, are perceived categorically. Musical intervals and scales, although they depend in part on acoustic factors, are primarily psychocultural entities—not mathematical or physical. Intervals are historically and culturally variable distances that are learned from oral traditions. There is no perfect tuning for any interval; even octaves are stretched relative to 2:1. Twelve-tone equal temperament is not intrinsically better or worse than Just or Pythagorean. Ratio theory is an important chapter in the history Western musical thought, but it is inconsistent with a modern evidence-based understanding of musical structure, perception and cognition.

Received: June 16, 2016, accepted September 14, 2017.

Key words: Pythagoras, interval, ratio, intonation, pitch

From a modern psychological viewpoint, the Pythagorean approach to understanding musical intervals has always been problematic because it ignores perception. In the 4th century BCE, Aristoxenus wrote three books entitled "Elements of Harmony," sometimes called "Harmonics" (in Latin: *Elementa Harmonica*), of which most contents have survived. In the second and third books, he argued on the basis of musical experience and intuition that the basic elements of musical structure—intervals, scales, tuning, melody—do not depend on arithmetic proportions, as the Pythagoreans claimed, but on what we today would call auditory psychology: processes of auditory perception, cognition, memory, and recall. To understand music, we have to perceive it. To understand the musical effect or function of an interval, we have to listen to it, not make abstract calculations. To understand how melodies work, we have to perceive, remember, and reproduce them. Musicians are not aware of ratios as they perform melodies. Interval sizes vary on a continuous scale and do not generally correspond to mathematically idealized ratios.

Neither Aristoxenus nor the Pythagoreans were in a position to test their ideas as we would today, in controlled experiments. It was not possible for them to measure interval sizes in real music performance with any accuracy. The empirical fields of research that we now take for granted to solve such problems (acoustics, psychology, psychoacoustics) did not exist in ancient Greece. It was not until the 19th century that scientists such as Helmholtz (1877/1954) and Stumpf (1883) developed the necessary technology, empirical methods, and theories of perception, which improved steadily during the 20th century. Since then, the findings of empirical research have been taken increasingly seriously—unfortunately, even when poorly carried out (Goldacre, 2010). But some philosophers have resisted, and Marxist philosopher György Lukács even declared empiricism to be "the ideology of the bourgeoisie" (Lukács, 1983, p. 174)—perhaps fearing that empiricism would prove him wrong.

Practical and epistemological issues of this kind can explain why the Pythagorean approach to understanding musical intervals persisted into the 19th century. But it is not so easy to explain why conceptualizations of musical intervals as ratios are still found today, in

Music Perception, volume 35, issue 4, pp. 475–501, issn 0730-7829, electronic issn 1533-8312. DOI: https://doi.org/10.1525/mp.2018.35.4.475

practically all areas of music research and scholarship: music acoustics (Caleon & Ramanathan, 2008; Fokker, 1949; Frosch, 1993; Hall, 1974), psychoacoustics of music (Mathews, Pierce, Reeves, & Roberts, 1988; Sethares, 1994), neuroscience of music (Bidelman & Krishnan; 2009; Foss, Altschuler, & James, 2007), music psychology (Cohen, Thorpe, & Trehub, 1987; Hannon, Soley, & Levine, 2011; Schellenberg & Trehub, 1996; Trainor, 1997; Trehub & Hannon, 2006), music biology (Araya-Salas, 2012; Gill & Purves, 2009; Schwartz, Howe, & Purves, 2003), music and mathematics (Honingh & Bod, 2005), the musical computer sciences (Cheung, 2013; Kirck, 1987; Schiemer, Alves, Taylor, & Havryliv, 2003), composition (Barlow, 1987; Gilmore, 1995; Hasegawa, 2006), performance (Fonville, 1991; Gratzki, 1993; Menke, 2009; Whitcomb, 2007; Williamson, 1939), music education (Dalby, 1992), ethnomusicology (Arnold, 1985), music history (Fauvel, Flood, & Wilson, 2003), music theory (Don, 2001; Vogel, 1955, 1975), and the history of music theory (Duffin, 2007; Godwin, 1992; James 1993).

In this article, we will present and explain the shortcomings of the Pythagorean ratio concept of musical interval and propose a psychocultural alternative. We will propose that the nature and origin of musical intervals—including their exact sizes—depend primarily on how they are perceived and experienced in real musical and cultural contexts, and only indirectly on the underlying physics. Music comprises complex patterns and structures that develop gradually over long historical periods in collective memory. To understand musical intervals ontologically, we must consider how they were perceived in different historical and cultural contexts, and how they were (re-)created (instantiated) and changed (adjusted) in those contexts.

Definition and Measurement Issues

"PYTHAGOREAN"

Little is known about Pythagoras himself, who lived in Greece in the late 6th century BCE. But his influence on later "Pythagorean" philosophers, including Plato, was considerable (Russell, 1945).

We will use the term "Pythagorean" in three different ways, all of which are commonly encountered. The first definition is more general and refers to the entire Pythagorean worldview, which is best understood in a broader cultural, historical, and philosophical context. In the 4th Century BCE, Plato theorized that the physical phenomena of everyday objects and human existence were an imperfect realization of an ideal perfect world lying behind them, and that human existence should endeavor to attain to that perfect world. A glimpse into that utopia was provided by the world of numbers, which Plato believed underlay every aspect of God's perfect world—including the harmony of the spheres, of which human music was an imperfect realization.

The second usage of "Pythagorean" is more specific. It refers to a musical tuning system based only on perfect octaves (P8s) and perfect fifths (P5s), tuned to ratios 2:1 and 3:2 respectively—as opposed to Just tuning, which also admits factors of 5. For example, a Pythagorean major third (M3) interval is 81:64, whereas a Just M3 is 5:4.

In a third usage, any ratio-based approach to understanding or performing musical intervals may be called "Pythagorean." This usage is more general than the second, because both Pythagorean and Just tuning are ratio-based. In this paper, we primarily argue against "Pythagoreanism" of this third kind. Regarding Pythagorean versus Just versions of a M3 (or any other interval), we are impartial: conceptually, both are equally (in)correct.

A consideration of the first usage of "Pythagorean" is beyond our scope. Before criticizing the third (which includes the second), we should ask to what extent it might still be correct. Of course the most important intervals in Western music and many other musical systems are P8, P5 and P4, in that order. Of course their tuning corresponds approximately to 2:1, 3:2, and 4:3 respectively, and these intervals can also be tuned more exactly (reliably) than other intervals (Vos, 1986). But the relationship between prevalence and ratio is indirect. Many sounds that we hear in everyday life, including voiced speech sounds, are harmonic complex tones. The distances between the partials *correspond to* ratios for physical reasons. But that does not necessarily mean they are *perceived as* ratios, such that the ratios are encoded in the corresponding neural mechanism or directly influence subjective experience. If intervals are not perceived that way, they cannot function in music that way.

To explain the concept of musical intervals, we prefer a parsimonious theoretical approach—one based on few assumptions. The relationship between intervals in music and intervals in non-musical sounds can be understood if we assume that the ear learns pitch distances from non-musical sounds and applies them to music, depending on their statistical distribution. The ear is exposed to a constant stream of pitch distances between both spectral components and fundamental frequencies, in both musical and non-musical contexts. The intervals between the harmonic partials of voiced speech sounds, and musical intervals, are two specific examples of this more general phenomenon. Since music is a social phenomenon, the term "the ear" also refers to groups of hearing individuals with a shared musical and auditory culture.

To apply non-musical intervals to music, the ear does not need to be sensitive to ratios as such. Evidence from modern intonation studies suggests that even if the ear were somehow directly sensitive to frequency ratios between everyday or musical sounds, the effect would be masked by other influences on intonation (expressive, music-structural, technical) and quasi-random variations.

INTONATION

Cambridge Dictionaries online (5/4/2016) define intonation as: 1) the sound changes produced by the rise and fall of the voice when speaking, especially when this has an effect on the meaning of what is said, e.g., "The end of a sentence that is not a question is usually marked by falling intonation," or 2) the degree to which the notes of a piece of music are played or sung exactly in tune, e.g., "The violinist had good intonation, and a wonderful pure tone."

The latter is close to our intended meaning, but lacks an operational definition of "in tune." For the purpose of this article, we will define intonation as the real-time adjustment of (fundamental) frequencies in music performance. That sounds straightforward, but to understand and apply this definition, some additional explanation is necessary.

Physical constraints of perception. According to the uncertainty principle in physics, the shorter the duration of a tone, the less exactly we can know its frequency. In music, tones must be extremely short (a few milliseconds, or a few periods) before their pitch becomes unclear or disappears (Houtsma, Rossing, & Wagenaars, 1987, Demonstration 13). The effective spectrum analysis window of the ear is adapted to optimize environmental interaction (Heinbach, 1988; Terhardt, 1998). The empirical literature suggests that the corresponding just-noticeable difference (JND) never falls below about 2-3 cents for the best expert listeners (Parncutt & Cohen, 1995). It follows that any discussion of tuning that considers fractions of a cent is psychologically implausible. This immediately counts out many music-theoretic writings based on ratio theory.

Inharmonicity. The term "fundamental frequency" suggests that musical notes are harmonic complex tones; that is, tones whose partial frequencies correspond exactly to a harmonic series. But that is not quite true for freely vibrating stings (piano or guitar); the inharmonicity of very low piano tones can change their pitch by as much as two semitones (Anderson & Strong, 2005). The word "fundamental" should therefore be excluded from a definition of intonation. But what should we replace it with? As the piano tone example illustrates, the frequency of the lowest partial does not necessarily correspond to the perceived pitch. That pitch might lie as much as a semitone away from the lowest partial if there is inharmonicity or pitch shifts, and it may be an octave away if there is pitch ambiguity. At this point, intonation becomes undefinable in purely physical terms.

Pitch versus frequency. In general, frequency is not the same as pitch. We use the term "pitch" in the psychoacoustic sense of the experience of how high or low a tone sounds. It is a purely subjective parameter—an experience of the listener that can depend on several different physical parameters. If we hear an oboe and a violin playing the same note, and the violin sounds slightly sharp relative to the oboe, we are making a relative pitch judgment. It is not necessarily true that the fundamental frequency of the violin is higher than the fundamental frequency of the oboe; to answer that question would require a physical measurement. To confirm that the violin sounds sharper in pitch (as opposed to sharper in timbre), we would need to determine the pitch of each tone separately by a standard method, such as comparing it with a pure reference tone of adjustable frequency and taking that frequency as a measure of pitch (Terhardt, 1972). An experiment in which listeners compare the violin and oboe sound directly will not work because listeners cannot generally separate pitch from timbre (Madsen & Geringer, 1981).

Complexity and subjectivity. Like pitch judgments, intonation judgments (Does an interval or chord sound in tune or out of tune?) are generally subjective; different listeners will experience the same sound differently, or offer different opinions. What we call "good intonation" is adjustment of frequencies in performance such that (most) (expert) listeners perceive them to be "in tune" in some sense—perhaps specific to a given style or corresponding to a composer's assumed intentions. The underlying criteria may be complex and opaque, and their relative importance may be a matter of style and taste.

Intonation, then, is the real-time adjustment of perceived pitch in music performance. This definition can only be understood and applied if we have standard, valid, reliable methods for measuring tone frequencies, pitches, and in-tuneness—regarded as three separate issues—and a good understanding of the complex relationship between physical measurements and subjective evaluation.

PITCH SHIFTS

The relationship between frequency and pitch has been the subject of innumerable empirical studies in psychoacoustics. The pitch of a pure tone, as measured by a standard procedure, deviates slightly and systematically from its frequency. The ultimate cause of such pitch shifts may be a nonlinearity in the cochlea's spectrum analysis, which is not surprising considering the enormous range of sound levels over which our hearing successfully operates. Seen another way, pitch shifts may be auditory illusions, analogous to optical illusions such as lines that are physically parallel but seem skewed (Terhardt, 1987).

The existence of pitch shifts immediately casts doubt on the Pythagorean idea that intervals are frequency ratios. If we could speak of pitch ratios (which we cannot, because pitch is not measured in hertz or cycles per second), the evidence suggests that they would often be perceptibly different from frequency ratios. That might be reasonable if we had an idealized cosmic concept of music, as the Pythagoreans did—something to do with the vibrations of the universe. In that case, only the frequencies would count. But the Pythagorean concept is not OK if we conceive of music as a human cultural product and activity.

Psychoacoustic research has revealed three kinds of deviation between a tone's pitch and its frequency:

Intensity. The pitch of a pure tone depends slightly on its physical intensity (Békésy, 1963; Burns, 1982; Egan & Meyer, 1950; Schubert, 1950; Stoll, 1985; Terhardt, 1974; Verschuure & van Meeteren, 1975; Webster, Miller, Thompson, & Davenport, 1952). The pitch of a low pure tone falls as it gets louder (Houtsma et al., 1987, demonstration 12, track 27-28). The effect is smaller for harmonic complex tones whose pitch is determined mainly by higher harmonics (Ritsma, 1967), which can explain why musicians seldom notice such pitch shifts in performance. The effect can nevertheless become audible in music composed of complex tones if loud music is heard quietly in the background. Consider loud rock music that is being played a long way away (a party down the street). At first one may only hear the bass line, and hence the key or tonality. When one later starts to hear a vocal or instrumental melody, it seems to be in a different key, as if the music were bitonal—suggesting a pitch shift of a semitone or more. To our knowledge, this particular effect has never been systematically studied; but there is a convincing demonstration of octave stretch in Houtsma et al. (1987, demonstration 16, track 32), which becomes a demonstration of the effect of loudness on the perceived size of an octave if it is played at different sound levels. Thus, the interval that we perceive between two tones depends generally on their relative loudness, even if the effect is often too small to be musically important.

Masking. The pitch of a pure tone is shifted by the introduction of another simultaneous tone or noise band (masker), the tone seeming to move away from the masker (Allanson & Schenkel, 1965; Hesse, 1987; Sonntag, 1983; Terhardt & Fastl, 1971; Webster & Schubert, 1954; Webster et al., 1952). Thus, the interval that we perceive between two simultaneous pure tones may be greater when the same two tones are presented simultaneously than when they are presented successively. More generally, the pitch of a pure tone (spectral pitch) depends on the spectrum in which it is embedded. Pitch shifts due to simultaneous sounds can approach a semitone, but the shifts of partials within a typical harmonic complex tone are no more than a few tens of cents (Terhardt, Stoll, & Seewann, 1982, Figure 7)—comparable with the theoretical difference between Just and Pythagorean tunings of a M3 (the syntonic comma, 81:80 or 22 cents). Pitch shifts of this kind can explain why a partial within a harmonic complex tone can be shifted by a quartertone or even a semitone and still be perceived as part of the tone (Moore, Glasberg, & Peters, 1986). While there are large individual differences in such pitch shifts, the direction of the shift is usually the same for all listeners (Stoll, 1985).

Diplacusis. A pure tone can evoke a different pitch in each ear, an effect known as binaural diplacusis. Burns (1982) also found that "pitch-intensity functions are often significantly different in the two ears of a given subject at a given frequency" (p. 1394).

The mainstream psychoacoustic literature of the past two decades has downplayed the importance of pitch shifts. Moore (1989, p. 166) remarked that the effect is small, generally amounting to less than 1% of frequency. This passage in his book was not revised in later editions (e.g., Moore, 2012, p. 243), reflecting a general reluctance (at least in the English-speaking research community) to investigate the phenomenon in more detail, in spite of its evident importance for a general understanding of pitch perception and for the quantitative testing and comparison of competing pitch theories. For a musician concerned with ensemble intonation, or a Pythagorean music theorist, 1% or 17 cents is a large shift; and as Moore noted, the shift gets larger for musically high or low frequencies, approaching perhaps 5% or almost a semitone.

Given the above three kinds of pitch shift and the body of research that confirms their existence, one can hardly deny that the pitch of a pure tone *generally* deviates from its frequency. Nor can one feign surprise when presented with experimental findings about the subjectivity and variability of preferred tunings, such as Rosner's (1999) finding that preferred tuning depends on pitch register. It follows that there cannot be a physically "perfect" tuning for any given interval, because if there were, our perception of it would change as sound levels changed, or as the interval was heard in different contexts, or depending on which ear we heard it with. This contradicts the Pythagorean concept, in which intervals sizes are held to be exact and constant.

INTONATION IN MUSICAL CONTEXTS

Empirical studies of intonation suffer from a measurement problem. The size of the interval between two selected tones (whether simultaneous or successive) in a music performance cannot, strictly speaking, be exactly measured, because each tone has finite duration (whereas the mathematical theory of spectrum analysis assumes infinite duration) and the frequency of each tone is usually constantly varying. A singer or cellist will adjust tuning during the course of a tone while monitoring the pitch of other instruments or singers. The frequency may rise and fall in a controlled vibrato, or it may vary quasirandomly. Such variations may be only partially under the performer's control; when interpreting empirical data, the extent to which tuning variations are deliberate is generally unclear.

Tone frequency measurements in recordings of music performances usually depend on a series of assumptions—often unstated. First, we assume that each tone in a performance contains a central, steady-state portion, whose start and end can be identified and whose frequency corresponds both to the performer's intention. Second, we assume that the perceived pitch of a tone corresponds to its frequency a certain time after the physical onset (say, 100 ms). Third, we assume that an intended or perceived pitch exists even if the frequency of the tone is oscillating in a regular way (vibrato). Fourth, we assume that the intended frequency of a scale step is independent of its context. Fifth, we assume that the performer achieves her intended pitch. All five assumptions are problematic.

1. The fundamental frequency of a musical tone typically changes throughout the tone, even if that change is imperceptible. Vocal jitter is an example; the degree of jitter rises as the sound level falls (Orlikoff & Kahane, 1991), so jitter may be a bigger problem when measuring intonation in quieter than louder vocal performances.
2. The intended pitch of a tone may be delayed for expressive purposes. Ascending fundamental frequency glides to target pitches are common in expressive singing; the underlying code may be borrowed from speech prosody (Sundberg, 1998). Tone languages are another case in which pitch targets are asymptotically approached:

   > Mandarin can be considered, under this framework, to have two static pitch targets—[high] and [low], and two dynamic targets—[rise] and [fall]. They are associated with the four lexical tones in Mandarin: H (high), L (low), R (rising) and F (falling), respectively . . . Assuming that the larynx, which implements the pitch targets, cannot change its state instantaneously, approaching a target always takes time, unless the target pitch is already reached before the onset of the host. As a result . . . there is often an apparent transition from the initial F0 at the onset of the host to the F0 contour later in the host that resembles the pitch target more closely. (Xu & Emily Wang, 2001, pp. 321-322)

3. If the frequency variation of a vibrato tone is roughly sinusoidal (a simple frequency modulation), the intended frequency corresponds to the mean fundamental frequency (Brown & Vaughn, 1994). But the greater the vibrato, the less clearly we hear the pitch, and the more tolerant we become for mistuning; vibrato may even be used to cover up poor intonation (cf. Prame, 1997).
4. The frequency of a scale step may depend on context. For example, a leading tone may be tuned differently in ascending and descending passages, as a signal to the listener that movement in a given direction is imminent (Devaney, Mandel, Ellis, & Fujinaga 2011; Fyk, 1995; Rakowski, 1990).
5. The pitches that performers intend in music performance may be different from what they actually produce. First, the pitch that a performer perceives when performing a tone may not be the same as the pitch perceived by the audience, especially if the performer experiences the tone as much louder (e.g., an opera singer). Second, there may be technical limitations. To measure intended pitches, Arom, Léothard, and Voisin (1997) presented

Central African musicians with an electronic xylophone, and Indonesian musicians with an electronic metallophone. In both cases, the musicians were invited to manually adjust the tuning and to check it while performing music from their repertoire. This was done repeatedly and the final results were interpreted as intended tunings. The advantage of this method is that it measures intentions independently of technical limitations. But it also has a disadvantage, namely the assumption that each scale step has a fixed target pitch that is independent of context. The truth may be that every pitch and every interval has a certain degree of uncertainty. We need to have a feel for that uncertainty when considering questions of intonation.

Variations in Interval Size in Performance

Psychological research has not been supportive of the Pythagorean concept of musical intervals as ratios. Experiments have revealed consistent stretching of musical octaves and other larger intervals in performance relative to simple ratios, and compression of smaller intervals (Rakowski, 1985). To our knowledge, no recording of a music performance by voices or instruments with real-time pitch adjustment (such as typical wind or string instruments) consistently conforms to either Just or Pythagorean intonation.

This raises an interesting question about the relationship between neurophysiological processes of pitch perception and musical intonation. The role of temporal processes (including the direct perception of periodicity) in virtual pitch perception has been thoroughly investigated (Cariani, 1999; Langner, 1997, Meddis & O'Mard, 1997; Patel & Balaban, 2001; Patterson, 1973, 1987). But we know of no evidence that processes of this kind affect the intonation of music in which performers can freely adjust their intonation from one moment to the next, such as most vocal, string, and wind music. The empirical literature on intonation in Western tonal music suggests instead that interval sizes are normally and unimodally distributed around mean interval sizes that do not differ significantly from familiar piano tuning—twelve equal divisions of the (stretched) octave. The literature also suggests that theoretical Just and Pythagorean variants lie well within those distributions.

Incidentally, we avoid the term "equal temperament" because it suggests that ratios are the norm from which other tunings are departures. On the assumption that no such standard exists, we prefer the term 12-EDO (twelve equal divisions of the octave), used today by musicians who compose and perform in different scales and tunings, as well as music researchers (e.g., Milne, Sethares, & Plamondon, 2008).

Why are musical intervals not tuned to ratios, although temporal information and phase-locking are maintained along neural pathways? A possible reason is that this temporal information does not affect conscious awareness when natural sounds are perceived, because monaural phase relationships carry no useful information for the perceiver. Most sounds heard in everyday environments are superpositions of direct and reflected sound, in which phase relations are jumbled (Parncutt, 2012; Terhardt, 1988).

THE PRACTICAL IMPOSSIBILITY OF JUST INTONATION

In Pythagorean tuning, every interval is a combination of P8s and P5s. It can therefore be expressed mathematically in the form 2^a3^b:1, where a and b are whole numbers. The Pythagorean system has the advantage that it allows us to define a single precise frequency for every tone that can be notated in conventional music notation: If we assume that A4 is 440 Hz, C♯5 is a major third (M3) above it, so its frequency in Pythagorean tuning is 440 Hz x (81/64) = 557 Hz—slightly sharper than 12-EDO, where C♯5 is 554 Hz. The Pythagorean frequency for D♭5 is slightly lower (if D♭ is considered to lie a M3 below F), but again clearly defined.

In Just tuning, interval ratios can include powers of 5 and are written in the form $2^a3^b5^c$:1, which allows for a M3 to be tuned as 4:5. But in Just tuning it is not possible to tune an entire diatonic scale without creating ambiguities. If for example the interval between scale degrees 1 and 2 of a major scale is 9:8 (a M2) and between 1 and 6 is 5:3 (a M6), the interval between 2 and 6 is not 3:2 (a P5) but 27:16. If one tries to preserve Just intervals in every sonority in a polyphonic texture, the frequencies of scale steps shift and the whole piece of music drifts up or down (Palisca, 1994, cited in Devaney, Mandel, & Fujinaga, 2012). But perceptible changes in the pitches of scale steps, from one sonority or passage to the next in a music performance, are hallmarks of poor intonation.

OCTAVE STRETCH

Perhaps the simplest claim of the Pythagoreans was that an octave has a ratio of 2:1. But even that is untrue—or at least misleading. When musicians play P8 intervals either harmonically or melodically, they are on average slightly wider than 2:1 (Burns & Ward, 1982; Corso, 1954; Kantorski, 1986; Loosen, 1993; Sundberg & Lindqvist, 1973). If two successive pure or complex tones are presented to a listener who then adjusts the interval to a P8, or if different tunings of a successive P8 are presented and the listener evaluates them, the preferred interval is typically stretched by 20 cents (Dobbins & Cuddy, 1982)—sometimes 50 (Walliser, 1969a). Carterette and Kendall (1994) also found stretched octaves in Javanese gamelan music, between non-harmonic complex tones.

How can preferences for stretched octaves be explained? There are two main theories, both of which are consistent with a psychohistorical, non-ratio concept of interval.

The piano. Octaves on the piano are physically stretched because the partials of a single tone are stretched relative to a harmonic series (Martin & Ward, 1961). Piano tuners tune octaves by avoiding beats between upper partials. The pervasiveness of the piano in 19th- and 20th-century Western musical culture may have made it a psychological standard for the tuning of other instruments.

Pitch shift. The spectral pitches of partials within typical environmental complex tones such as voiced speech sounds are shifted due to the simultaneous presence of other tones. Since spectral pitches are defined as purely subjective experiences, the effect is entirely psychological. The auditory system becomes familiar with a stretched harmonic series and then "expects" musical intervals to be similarly stretched (Terhardt, 1979). Individuals may learn octave stretch either from the spectral pitch patterns of single harmonic tones or from music in which octaves are stretched.

The expression "stretched octave" is misleading, because it suggests that the "real" octave is the ratio 2:1. In fact, the "real" octave is the subjective octave as played in music and experienced by music listeners—the interval at which men and women/children have sung for millennia. It may be more correct to say that the ratio 2:1 is smaller than this subjective octave. The idea of a *physical octave* or *mathematical octave* (exactly 2:1) can nevertheless be useful. It is a *reliable* definition of an octave, because it always has exactly the same physical size. But if we ask musicians to tune octaves by ear, the result is a more (ecologically) *valid* measure of an octave, especially if this is done in the context of a musical performance.

COMPRESSION OF SMALLER INTERVALS

Rakowski (1976) asked trained music students to tune musical intervals between successive pure tones. One tone was constant at 125, 250, 500, or 1000 Hz; the other could be freely varied. The students consistently tuned smaller intervals smaller than 12-EDO and larger intervals larger. Rakowski (1985) reported a similar result using both pure and complex tones and a reference tone of 500 Hz; intervals greater than seven semitones were stretched (P8s by about 20 cents on average) and intervals smaller than five semitones were compressed (the m2 by about 10 cents). Rakowski's results did not depend on spectral envelope (waveforms were sinusoidal, triangular or square), suggesting that interval sizes were learned from pitch patterns in music.

Rakowski (1994) repeated this experiment using a wider range of intervals and confirmed that intervals of an octave or more are consistently stretched. Similarly, Vurma and Ross (2006) asked professional singers to sing m2, tritone, and P5 intervals; m2s tended to be sung smaller than 12-EDO and P5s wider (both rising and falling). In a listening experiment, Rosner (1999) confirmed that musicians prefer larger intervals stretched and smaller intervals compressed, and additionally showed that the size of stretching and compression depends on pitch register—again consistent with a psychocultural concept.

How can this difference between small and large intervals be explained? In Bregman's (1990) theory of auditory scene analysis, larger melodic intervals (perhaps fourths and larger) tend to break up a melody in the listener's imagination: tones spanning such intervals are perceived as belonging to different melodies (streams). Tones spanning smaller intervals are perceived to belong to the same melody and promote melodic coherence (cf. Huron, 2001). In the theory of melodic expectation of Narmour (1992), melodic intervals greater than a P5 are considered "large" and imply a consequent reversal of registral direction. These musical effects may have a foundation in non-musical environmental interaction: frequencies that are closer to each other are more likely to originate from the same sound source and so to fuse into one perceptual object in the listener's experience.

The tendency to tune larger intervals too large and smaller intervals too small may be an instance of a more general tendency to exaggerate differences in performance for expressive purposes (Parncutt, 2003; Sundberg, Askenfelt, & Frydén, 1983). Performers may exaggerate the size of larger intervals to clarify the separation of different melodic streams, and reduce the size of smaller intervals to enhance coherence within streams.

Vos and Troost (1989) observed that the most common interval between successive tones in melodies from different cultures is approximately one whole tone or two semitones, corresponding to the music-theoretical concept "scale step." Miller and Heise (1950) demonstrated that the sensation of "trill" in music disappears (it turns into a "shake") when interval size exceeds a threshold of about two semitones, and Shonle and Horan (1976) found the "trill threshold" to be about a quarter of a critical band or about one semitone in the central musical range. Thus, a trill sounds like a variation in the pitch of one tone and works only for intervallic steps, whereas a shake sounds like several separate tones being repeated and applies to intervallic leaps. Intonation studies point to a general tendency to reduce the size of such small intervals.

Pianists cannot adjust intonation in this way, so if one were to claim that a well-tuned piano sounded out of tune, that would be a possible reason. Unlike a reason based on ratios, this one would have an empirical foundation. But one could also argue that pianists have a tuning advantage over other instrumentalists, because every scale degree on the piano has a fixed frequency, which by itself can give an impression of good intonation. Violinists and flutists strive to maintain fixed frequencies for scale steps, but may deviate from this ideal when reducing the size of small intervals. In this and other ways, good tuning often involves compromises.

THE ABSENCE OF CONSISTENTLY JUST OR PYTHAGOREAN INTONATION IN REAL PERFORMANCE

A Just M3 interval is 5:4 or 386 cents; the Pythagorean variant is 81:64 or 408 cents. If intervals were ratios, we would expect performers to consistently favor one or the other of these. If we measured many M3s in performance, we would expect to find a normal distribution with a peak at either 386 or 408 cents, or a bimodal distribution with both peaks.

To our knowledge, no such result has ever been reported. Observed tendencies toward Just and Pythagorean may merely be tendencies toward intervals that are slightly larger or smaller than 12-EDO. Many studies have found that the M3 interval in the best performances lies near to 12-EDO (400 cents) or between 12-EDO and Pythagorean (Duke, 1985; Karrick, 1998; Loosen, 1993; Mason, 1960; Morrison, 2000; O'Keefe, 1975; Sundberg, 1982). A tendency toward Pythagorean tuning for string players under ideal conditions (Greene, 1937; Nickerson, 1949) may be due either to the P5 intervals between the open violin strings, which players tune before playing, or the leading tone effect of Fyk (1995). Another way to investigate musical intervals psychologically is to present the same music in different tunings and ask listeners to evaluate the tuning or to say which version they prefer; tunings that lie between 12-EDO and Pythagorean tend to be preferred (Loosen, 1995).

Roberts and Mathews (1984) studied intonation preferences of listeners presented with chords of steady-state harmonic complex tones, and reported that "One group preferred chords in just intonation and their preferences decreased monotonically as the intonation deviated from just intonation; the other group preferred intonations that deviated from just intonation by ± 15 cents" (p. 952). In other words, some preferred something approaching 12-EDO or 400 cents (which is what most empirical studies have found) and others preferred intervals with a minimum of beating. Presumably, the tones in this experiment had relatively long durations, attracting listeners' attention to beats. In most musical contexts, such beats are inaudible, which can explain why this finding has not been replicated. Hagerman and Sundberg (1980) found that barbershop quartets often preferred intonations that, on average, were on the Just side of 12-EDO, but they also found that slight mistunings relative to Just did not produce audible beats. Another example of beats that are audible under ideal experimental conditions but almost never in music is beating between pure tones spanning mistuned intervals in isolated sounds (Plomp, 1967). A composer might use such effects in music that is intended for headphones only.

The empirical literature suggests that tendencies toward Just tuning in musical chords are confined to relatively slow music in which the tones are sustained without vibrato and are of roughly equal amplitude. If that is true, it can be explained in two ways. First, the auditory system may be familiar with the distance between harmonics 4 and 5 in environmental harmonic complex tones; musicians may try to imitate this interval when performing M3 intervals, and listeners may simply like it because it is familiar (albeit unconsciously) from voiced sounds in speech. A second explanation is that we prefer chords in which harmonics of one tone overlap with harmonics of another to reduce beating. If three conditions are fulfilled—the tones are exactly harmonic with little or no vibrato, the music is performed at a slow tempo, and nearby partials have almost the same amplitude—Just intonation best solves the beating problem (O'Keefe, 1975). These two different processes lead to a similar musical result, so if something approaching Just tuning is observed, it is not clear which process is responsible. In any case it is not the ratio itself that is causing the phenomenon—it is familiarity with the harmonic series and/or a culture-specific (or non-universal) aversion to beats and roughness, as opposed to preferences for rough seconds in Lithuanian Sutartines (Ambrazevičius & Wiśniewska, 2009).

There are also different ways to explain tendencies toward Pythagorean intonation. One is that leading tones are raised in anticipation of their resolution, or the difference between M3s and m3s is exaggerated to reduce the chance of confusing them (Fyk, 1995). Another is that a general preference for exactly tuned P5 intervals can determine the tuning of musical scales, which in turn determines the tuning of individual intervals, including the M3 or M7.

Historically or prehistorically, something approaching Pythagorean intonation may have emerged when scales were constructed from chains of rising P5 and falling P4 intervals (or vice-versa). The standard pentatonic scale represented by the black keys of the piano (in any transposition) can be expanded to the white-note diatonic by adding tones that are a P5 above or below existing pitches: starting with the pentatonic set CDEGA, F may be added because it is a P5 below C, and B because it is a P5 above E, creating the diatonic set CDEFGAB. A possible prehistoric scenario is a gradual evolution of scale structures as listeners become sensitive to the tuning of P5 intervals, so that intervals in the vicinity of a P5 gradually approach it. Changes of this kind might happen when melodies pass from one generation to the next in an oral tradition (Parncutt, 2001). The process may involve avoiding beats when tuning instruments, leading to more exact tuning (Wolfe & Schubert, 2008). Such processes may explain why, in ancient Greece, the mysterious enharmonic genus, in which the P4 interval was divided into a M3 and two quarter-tones, gradually became less common, and the diatonic (M2+M2+m2) and chromatic genera (m3+m2+m2) became more common (Chalmers, 1993; Mathiesen, 1999).

INTONATION OF SOLO VERSUS ACCOMPANIMENT

Further evidence against the idea that intervals are ratios is the observation that soloists tend to play or sing sharp relative to their accompaniment (Kantorski, 1986). Barbour (1938, p. 55) noted that "In fact, two writers have emphasized sharp singing as a general characteristic—and sharp singing is incompatible with just intonation" (the writers being Schoen, 1926, and Cameron, 1907; both references cited in Barbour, 1938).

Why would a soloist play or sing sharp? One motivation may be to help the solo line stand out against the accompaniment (a kind of streaming effect; Bregman, 1990). Another reason may involve timbre, the goal being a kind of solistic brilliance. A third explanation is simply that soloists lead while others follow; soloists may lead both in time (melody lead: Rasch, 1979) and in pitch (by playing sharp).

Yet another possible explanation for sharp solo intonation involves pitch shifts. The tones may be physically sharp, but not sound sharp. The pitch of a pure tone goes slightly flat as the tone becomes louder; this effect also becomes stronger the further one departs from about 2 kHz (Terhardt, 1974). If a solo musician were playing pure tones, their pitch would go flat as they got louder. The soloist would therefore increase their frequency to get them back in tune. This explanation is problematic because soloists do not play pure tones—they play complex tones whose upper partials tend to be strong relative to the fundamental (so their timbre is relatively bright): in general, when you play an instrument louder, the timbre becomes brighter. However, the pitch of complex tones in the central musical range is mainly determined by higher harmonics (Ritsma, 1967), which explains why pitch shifts with changes of loudness have not been observed for harmonic complex tones (Stoll, 1985).

THE INHERENTLY APPROXIMATE AND UNCERTAIN NATURE OF INTONATION

The empirical research on intonation in music performance has consistently revealed a high degree of intonational uncertainty, even in the best performances by the best performers, or in performances that listeners and performers agree have excellent intonation. Vurma and Ross (2006) asked individual professional singers to sing m2, tritone, and P5 intervals; they recorded their performances and played them back to the same group, asking them to evaluate the intonation. The authors concluded that such melodic intervals can be 20-25 cents out of tune and still be judged to be in tune by expert listeners. Delviniotis, Kouroupetroglou, and Theodoridis (2008) asked expert singers of Byzantine Chant to sing ascending and descending scales; the authors reported a standard deviation of 30 cents for intervals within these scales. Devaney et al. (2012) asked professional and semi-professional vocal ensembles who specialize in Renaissance polyphony to sing simple three-part chord progressions slowly under ideal conditions; careful measurements revealed standard deviations of typically 10 ± 3 cents for ascending and descending whole-tone intervals and for the vertical intervals m3, M3, P4, P5, M6, and P8 (see also Prame, 1997). Listeners in Devaney's experiments usually did not notice any mistuning.

If that is true, how can we explain continuing beliefs in the use of Just intonation by the best vocal ensembles when singing Renaissance polyphony (Duffin, 2007; Havrøy, 2013)? The beauty of their timbre suggests that they must be using a special kind of intonation, and from a theoretical viewpoint Just intonation seems like the obvious candidate. In fact, those ensembles may merely be singing close to 12-EDO with very steady tones (not slides or vibrato) and a characteristic timbre that listeners associate with this style of music.

There are many reasons to reject the ideal of Just intonation for Renaissance polyphony or any other music. First, a Just diatonic scale in which all intervals between all pairs of tones are Just is a mathematical impossibility; some kind of adjustment or temperament is always necessary. It is true that singers in performance are constantly adjusting their intonation to approach optimal interval sizes relative to the musical context, but we have no evidence that those optimal interval sizes correspond to Just ratios. Second, the tendency toward "pure" P5s, and the greater sensitivity of musicians to mistuning in P5s than in M3s and m3s (Vos & Vianen, 1985), means that tendencies toward Pythagorean intonation are possible and likely in any diatonically based music—but that does not mean Pythagorean tuning is ideal, either. Third, all modern performers are influenced by their experience of different kinds of tuning, including the tuning of the piano (Kopiez, 2003). For any one or more of these reasons, attempts to approach Just intonation tend to sound out of tune.

Intonational accuracy, and hence auditory tolerance for mistuning, depends strongly on context. Rakowski (1990) had musicians tune musical intervals in isolation and in a musical context, and found that the range of acceptable tunings was smaller in a musical context. Presumably, musical context gives a musician more information upon which to base frequency adjustment. On this basis, we might predict that the standard deviation of 10 cents in the measurements of Devaney et al. (2012) would have been higher if intervals had been measured in isolation rather than in a musical context. If intervals were essentially number ratios, as Pythagoreans suppose, one might expect the opposite: tunings of isolated intervals would be more accurate, corresponding to small integer frequency ratios as directly perceived by the brain. The presence of a musical context that by necessity departs from integer ratios generally gives performers conflicting pitch references, so we might expect tunings to deteriorate. That this prediction contradicts observations is another argument that intervals are distances and not ratios.

Performers of string and wind instruments may be expected to achieve higher degrees of intonational precision than singers. That may especially motivate them to consider theories of intonation based on ratios (Heman, 1964; Mantel, 2005). But the empirical literature suggests that that the intonation of individual tones, played by instruments with continuously adjustable tuning, in music whose intonation is judged good or excellent, typically varies by at least ± 10 cents—comparable to the syntonic comma (22 cents), the difference between the Just and Pythagorean M3. An uncertainty as small as ± 10 cents is only reasonably possible when tones have long duration and almost equal amplitude. The interval(s) should be consonant, and there should be no vibrato or expression. At faster tempos, intonation may vary by ± 50 cents, even in the best performances—even without vibrato, expressive variation of intonation, or expressive dissonance (Burns, 1999). In some styles, intonational variation can exceed a semitone, even in the best performances (e.g., romantic opera; Prame, 1997). Approaches to teaching performance based on ratios may attract attention to fine differences in intonation and motivate musicians to devote time and effort to this task, which ultimately improves their intonational skills. That does not however mean that the ratios are "correct."

The uncertainty of interval sizes may depend on music-historical processes whereby the tuning of P8s, P5s, and P4s between non scale-steps gradually changed, perhaps to enable different instruments to play together. If Pythagorean tuning were the ultimate aim of such a process, mistunings would accumulate across P8s and P5s (for example, the three P5s and one P8 that combine to produce a M6) in such a way that the resultant mistuning might become perceptible (bigger than a JND) even if the component mistunings were not (smaller than a JND). Consider the modern piano (with 12 equal divisions of a stretched octave): in slow consonant chord progressions, its M3 interval may be perceptibly different from Just and Pythagorean variants, even if the mistuning of the P5 relative to 3:2 is imperceptible.

The inherent uncertainty of intervals sizes is consistent with the theory of Helmholtz (1877/1954), who explained that frequency ratios arise from overlapping harmonic partials in simultaneous harmonic complex tones. To avoid perceptual roughness, the overlap need not be exact. Similarly, in Stumpf's (1883) theory it is also not necessary for overlap to be exact, if the aim is to produce the effect called perceptual fusion (*Verschmelzung*) that he assumed to be the foundation of musical consonance.

There have been several attempts to create spectrally or temporally based models of consonance and dissonance (e.g., Aures, 1985; Ayers, Aeschbach, & Walker, 1980; Hutchinson & Knopoff, 1978). Parncutt (1989) and Collins, Tillmann, Barrett, Delbe, and Janata (2014) termed these models "sensory" as opposed to more "cultural" or "cognitive" approaches. But when the predictions of sensory models are compared with human data on the perceived consonance/dissonance or frequency of occurrence of musical chords comprising harmonic-complex tones, a model based on notated intervals between fundamental frequencies performs as well as or better than a model based on the amplitude spectrum (Bigand, Parncutt, & Lerdahl, 1996; Parncutt, Reisinger, Fuchs, & Kaiser, submitted). A simple model involves counting the semitones (or the semitones and tritones) in a chord; another option is to count all intervals between all notes and weight them appropriately (cf. Huron, 1994). Considering the undeniable role of both (psycho-) acoustics and cultural learning, we may assert that consonance/dissonance (like other psychological percepts) depends generally on a mixture of physics/physiology (nature) and learning (nurture).

Given this background, when a respected theorist such as Ehrlich (1998) writes at length about the difference between a P5 of 708.8143 cents and a P5 of 710.0927 cents, we can be sure that something is wrong. If the frequency JND of two successive tones for trained listeners under ideal laboratory conditions is a few cents, it is misleading to cite musical interval sizes with a precision of less than one cent. In performance, tuning is limited by both human technical and auditory abilities. Moreover, the existence of categorical perception in all senses including interval perception suggests that the inherently approximate nature of musical intervals arises from the circumstances and processes of their perception and cognition.

A Psychocultural Theory of Musical Intonation

If musical intervals are not ratios, what are they? In a psychocultural theory of musical intervals and intonation, intervals are assumed to be both fundamentally psychological and fundamentally cultural in nature.

NEURAL FOUNDATIONS

In the 18th century, Euler and Leibniz believed that the brain could comprehend number ratios built from the first three prime numbers 2, 3, and 5 but not 7 or 11, which would explain the apparent absence of ratios including the prime numbers 7 and 11 in tonal music. They proposed that the human soul is not fine or subtle enough to perceive more complex ratios (Leissinger, 1994).

How might the brain perceive ratios when the distances between fundamental frequencies in music (assuming the tones are periodic or have harmonic spectra) vary on a continuous scale from zero to several octaves? Even if there are peaks in the distribution near the 12 equal-tempered intervals, it is still hard to imagine how the brain could directly perceive ratios in complex musical contexts.

We assume instead that neural foundation of musical intervals and their tuning and intonation is the biological neural network, which allows for context-dependent learning of interval sizes in both everyday physical and social environments as well as music (Koelsch & Siebel, 2005; Laden & Keefe, 1989; Patel, 2003; Tillmann, Bharucha, & Bigand, 2000; Todd & Loy, 1991; Zatorre, Evans, & Meyer, 1994). The ultimate neural foundation for the sizes and perceived qualities of musical intervals may be Hebbian learning and connectionism. The connectionist learning of musical intervals is a physically complex process that involves both temporal and spectral aspects and cannot be reduced to ratios or usefully explained in terms of ratios.

INTONATION AS COMPROMISE

The diversity of psychoacoustic, psychological, and music-structural factors that influence tuning and intonation suggest that there are no ideal solutions. Every intonation is a compromise among partially contradicting criteria (Terhardt, 1976). For example, the M3 interval is often performed somewhere between Just (386 cents) and Pythagorean (408). The piano's M3 interval is a compromise, but so is every other well-known tuning system. In this sense, the piano is not "out of tune"; on the contrary, it represents a useful tuning compromise for western tonal music. This statement is valid even for music performed or composed long before the invention of 12-EDO.

Our perception of good versus bad tuning may be based on either familiarity with intervals in music (memory) or spectral structures and relationships (real-time direct perception). We may prefer interval sizes that represent an optimal compromise between the psychoacoustic criteria underlying Just and Pythagorean tuning. Since explanations of this kind do not contradict each other, and good evidence and arguments can be found for all of them, they may all be correct. Each is sufficient, from a psychological viewpoint, to explain modern measurements of interval size in music performance. Taken together, they undermine the Pythagorean ratio concept.

The criteria that determine intonation in real music include the following:

- Minimizing beats and roughness within individual sonorities (favoring Just intonation),
- Approximating the interval sizes heard among spectral pitches in harmonic complex tones (again tending toward Just),
- Accurately tuning P8 and P5 intervals between scale steps (favoring Pythagorean),
- Assigning a single fixed frequency to every note in a diatonic scale, or an entire musical score (favoring Pythagorean or 12-EDO, since this is not possible in Just),

- Accommodating stretching of all intervals larger than about a P4,
- Enabling unlimited modulation to different chords and keys in the 12-tone chromatic scale (favoring 12-EDO), and
- Approaching musically familiar interval sizes.

It is often supposed that unlimited modulation is the main reason why 12-EDO is so widely accepted today (e.g., Barbour, 2004). But that does not necessarily explain why choirs and string or wind ensembles perform close to 12-EDO. Unlimited modulation is just one point in a list of arguments that point to slightly stretched 12-EDO as a good compromise solution. The relative importance of unlimited modulation also depends on style and genre.

In addition to these factors, sensitivity to mistuning may depend on the complexity of the musical context. To test this idea, we might envisage an empirical study to compare standard deviations in interval sizes within performed passages, comparing relatively consonant versus dissonant harmonies (e.g., major and minor triads versus bitonal bebop chords) and relatively tonal versus atonal contexts.

AUDITORY SCENE ANALYSIS AND ECOLOGICAL PSYCHOLOGY

Since intonation is about quasi-exact tunings of frequencies, we may ask why exactitude is so valued or desirable. Why are humans so sensitive to such small frequency differences?

According to Bregman's (1990) theory of auditory scene analysis, we recognize sound sources by recognizing patterns in frequency-time space. If the pitch of a speaking voice goes up and down during an utterance, the pitches of the audible harmonics go up and down with it (coherent frequency modulation). This characteristic pattern (which looks like parallel curved lines on a graph of pitch against time) is recognized by the auditory system, allowing a listener to track a speaking voice in a noisy background (cocktail party effect). The psychological reality of the effect is disputed: McAdams (1989) presented evidence in favor, Carlyon (1991) against.

This pattern-recognition process relies on a kind of running spectrum analysis of incoming sounds in the auditory periphery. Without this initial analysis, it would be difficult to recognize sound sources reliably. To understand why this is so, consider the physics of human acoustic interaction with everyday environments. The auditory world includes multiple acoustic reflectors: the ground, the walls and roof of a room, or any nearby object from which sound can be reflected. Just about every sound we hear at each ear is a superposition of direct and indirect (reflected) sound. The phase difference between the two depends on the path difference, which is usually much bigger than the wavelength. That makes the monaural phase difference between direct and indirect sound effectively random. The superposition of similar waveforms with different phases radically changes the shape of the waveform—so much that it is not generally possible to recognize (distinguish and identify) sound sources reliably on that basis.

Terhardt (1988, p. 6) explained as follows (translation RP):

> Only the frequencies survive transport through the linear system without noteworthy changes. The fact that essential musical information, namely pitch, is conveyed almost exclusively by the frequencies of the partials, means that their propagational resilience (*Übertragungsresistenz*) is an indispensable foundation for any enjoyment of music and evaluation of sound, whether it be in the concert hall, on a recording, or on the radio. We experience this state of affairs and take it for granted every day, but that does not make it any less remarkable from a scientific point of view; it can be regarded as a key to a deeper understanding of the process of musical (and incidentally also linguistic) communication . . . The *amplitudes* of the partials are much less reliable carriers of information. They are changed by the frequency response of the acoustic path between source and ear (*Übertragungsstrecke*) . . . For this reason, one can only make a very rough estimate of the change in amplitude of a given partial during transmission; the amplitude at the eardrum of the hearer can be several times more or less than the estimate. This means that most of the information carried by the amplitudes of individual partials is lost . . . To a greater or lesser degree, this quasi-random variation of partial amplitudes affects all sound features that depend on the detailed structure of the partial amplitude distribution, such as timbre, beating and roughness. The result for the phases of the partials is even more pessimistic. They do not transmit any useful information at all, because phase changes during transmission are almost entirely unpredictable.

This is an ecological-evolutionary explanation for the important role of auditory spectrum analysis in all animal hearing. An animal is more likely to survive and reproduce if it can interact successfully with its physical and social environments. How it does that depends on the affordances of environmental organisms objects—their action possibilities, that is, what they can be used for, dependent on both their physical properties and the user's capabilities (Gibson, 1979). That is why auditory spectrum analysis is common among animals of all kinds, including insects (Fonseca, Münch, & Henning, 2000). According to Reed (1996),

> The fundamental hypothesis of ecological psychology . . . is that *affordances and only the relative availability (or nonavailability) of affordances create selection pressure on animals; hence, behavior is regulated with respect to the affordances of the environment for a given animal* (p. 18; italics in original).

Frequencies are not changed by reflection and superposition, so animals can rely on them to provide information about sound sources. Amplitudes are changed considerably, but they still provide some useful information. Monaural phase relations between partials are usually jumbled so much that they might as well be random. The auditory system cannot entirely ignore them, because neural processes of pitch perception are fundamentally time-domain processes (Cariani, 1999; Langner, 1997; Patterson, 1987). That can explain why monaural phase differences between partials can shift the pitch of artificial sounds (Licklider, 1957) but not natural or quasi-natural sounds (Patterson, 1973); they nevertheless affect timbre (Plomp & Steeneken, 1969).

After auditory spectrum analysis, we recognize sound sources primarily from their characteristic frequency-time patterns (Bregman, 1990). Spectrum analysis enables us to separate signal from noise; in music, it allows us to focus on a soloist and ignore the accompaniment. The amplitude spectrum of a sound provides further clues to the identity or state of the source, but timbre also depends crucially on the *temporal* amplitude envelope (Houtsma et al., 1987, demonstration 29, track 54-56), consistent with the idea that the frequencies of the partials are more important for source recognition (and hence for music) than their amplitudes.

LEARNING MUSICAL INTERVALS

Loosen (1995) asked musicians to evaluate the tuning of simple rising and falling scales that had been tuned to Just, Pythagorean, and 12-EDO. Violinists preferred something approaching Pythagorean, pianists preferred 12-EDO, and nonmusicians had no preference. That suggests that tuning preferences are primarily learned—not primarily determined by the acoustics of the sounds themselves, or by the physiology of hearing. Learned intonation systems are remarkably robust; a professional trumpeter may play close to 12-EDO, even if the accompaniment is in Just intonation (Kopiez, 2003).

If musical intervals are learned, where and when are they learned, and from what source? If we consider the kinds of sounds to which humans are typically exposed in everyday life, we can identify two main environmental sources of information about musical intervals: intervals between audible harmonics within voiced speech sounds, and intervals between tones in music.

Terhardt (1972, 1976) assumed that the auditory system learns the pattern of intervals in the harmonic series by exposure to voiced speech sounds (vowels and voiced consonants)—a prerequisite for speech acquisition. The process presumably begins in early life, given the importance of fundamental frequency tracking for the prosodic meaning of speech (Pell, 2006).

The fetus can hear the mother's voice for 20 weeks before birth (Bibas et al., 2008; Pujol & Lavigne-Rebillard, 1995). During this time it is probably acquiring information about both the harmonic series and speech prosody. Ultimately, this "knowledge" may serve to improve the bonding relationship with the mother after birth and in that way enhance the probability of survival in ancient physical and social settings (Hepper & Shahidullah, 1994; Parncutt, 2009). This process also gives the infant a head-start on the long, complex, and socially essential process of speech acquisition. Empirical evidence for the sensitivity of the fetus to maternal speech (DeCasper & Spence, 1986) suggests the fetus is acquiring information of this kind throughout the third trimester. The fetus can hear fewer harmonics than the infant in voiced speech sounds due to low-pass filtering by the amniotic fluid, but frequencies between about 100 and 1000 Hz are often audible (Richards, Frentzen, Gerhardt, McCann, & Abrams, 1992).

These findings suggest that "knowledge" of the intervals between the lower harmonics of harmonic complex tones is present at birth and becomes more accurate and more complex in early infancy, given the ability of infants to track missing fundamental frequencies. He and Trainor (2009) found that infants make significant progress in this task between 3 and 4 months postnatal age.

These ideas have interesting implications for the tuning of the M3 interval. The Just version of this interval corresponds to the perceived distance between the 4th and 5th harmonics in a speech sound, in the absence of pitch shifts. If musicians or listeners are found in an experiment to prefer the Just M3, familiarity with the pitches of the 4th and 5th harmonics may be the reason. Another possible reason: when harmonic complex tones span this interval, some of the upper partials overlap exactly; the 5th (and 10th) harmonic of the lower tone coincides with the 4th (and 8th) harmonic of the higher. This reduces the audible beating between almost-coincident partials, which may be a perceptible in slow music composed of precisely tuned harmonic complex tones without vibrato. In this second case, the optimum interval size would be learned from music in a two-stage process: first, musicians approach the ratio 5:4 to avoid beats, and second, listeners become familiar with M3s of this size and subsequently prefer them. Note that in both cases the listener does not perceive the 5:4 ratio per se; any apparent sensitivity to this ratio is a consequence of other sensitivity to other phenomena.

Missing fundamentals are perceived by cats (Heffner & Whitfield, 1976), fish (Fay, 1984), and songbirds (Cynx & Shapiro, 1986). The underlying process evidently involves both the time domain (periodicity detection) and the frequency domain (recognition of incomplete harmonic series), just as it does in humans (Cheveigné, 2003; Moore, 2012). Presumably, these animals learn pitch intervals from complex tones in the environment just as humans do; there is no reason to suppose that human and non-human animals would be different in this respect. Preisler and Schmidt (1998) tentatively demonstrated ultrasonic missing fundamental perception in bats; the limited firing rate of auditory neurons points to a frequency-domain explanation involving recognition of harmonic patterns of spectral pitches. On this basis, one may ask why non-human species are not making music with intervals corresponding to distances between harmonic partials. In fact, they sometimes are: certain songbirds produce harmonic intervals more often than non-harmonic intervals (Doolittle & Brumm, 2012). But that does not necessarily make their songs "music"; the difference evidently involves human reflective consciousness.

THE REAL-TIME PROCESS OF ADJUSTING INTONATION IN PERFORMANCE

"[I]ntonation is an amalgam of several sub-skills including pitch discrimination, pitch matching, and instrument tuning" (Morrison & Fyk, 2002, p. 183). What are the underlying processes as performers adjust interval sizes in real time?

Morrison and Fyk (2002) continued: "[I]ntonation is a process that demands at least three levels of awareness—the performers' own actions, the actions of those around them, and an abstract model of ideal performance" (p. 192). To understand this process, we must separate perception of isolated intervals from perception of pitch relative to the prevailing scale. On the one hand, performers are constantly evaluating and adjusting the sizes of intervals between simultaneous and successive tones. On the other hand, they are adjusting the pitch of each tone relative to the prevailing scale that has been established in previous minutes of performance, and confirmed in previous seconds, even if this scale is out of tune compared to some intended standard such as A4 = 440 Hz. Drifts away from this standard may happen even in the best performances and go unnoticed by both performers and listeners.

Let us assume that a given interval (such as M3) has a perceptually ideal target size that is independent of context. This target size has been learned from experience of instruments such as the piano. In the real-time context of an ensemble performance, a performer such as a singer or violinist who is aiming to produce this interval must also deal with small shifts in the pitches of scale steps, as performers in the ensemble adjust their exact pitches according to different and sometimes conflicting constraints of intonation. Performers must respond to these shifts if good intonation is to be maintained; indeed that may be considered *the* central intonational skill. At any moment in time, a performer who is about to perform a tone does that in a context of slightly mistuned scale-step pitches produced by the ensemble. She must navigate within a web of interval relations between the tone being played and other structurally or texturally important tones, all of which may be slightly mistuned. Performers are constantly making compromises in their tuning to adjust for such differences, and these adjustments happen largely automatically—not unlike the tiny adjustments that a downhill skier constantly makes to the relative pressure on the two feet, direction of the skis, weight forward or backward, bending of limbs, and so on.

Effects of this kind might explain why some performers have the impression that the piano is out of tune. Explanations for this impression that are based on ratio theory are logically fallacious. There are several possible reasons why a piano accompaniment might sound out of tune to a solo violinist, and deviation from "ideal" ratios is just one of them. Performers are constantly compensating for pitch shifts on stage, adjusting the physical size of an interval to compensate for such shifts, so that the interval approaches its ideal subjective size. Thus, physical tunings *generally* deviate from "ideal" tunings, however defined.

Consider the case of medieval chant, as performed both then and now. In music theory, the ideal intonation of chant is Pythagorean (e.g., Bower, 2002; Krahenbuehl & Schmidt, 1962). Today, we can only guess how chant was intoned in the Middle Ages. Even then, intonation could have been closer, on average, to modern 12-EDO than Pythagorean, because even then 12-EDO represented a compromise between Just and Pythagorean, even if the idea of 12-EDO was unthinkable for the theorists of the time (let alone the singers).

Medieval listeners presumably perceived the spectral pitches at the 4th and 5th harmonics of an isolated complex tone, just as we do today (e.g., in overtone singing), which could have steered the intonation of chant in the direction of Just. Evidence for this can be found in statistical data on the prevalence of individual scale-steps in chant. In an unpublished computer-based analysis (similar to Parncutt & Prem, 2008) of the *Liber Usualis* using the database DDMAL (Thompson, Hankinson, & Fujinaga, 2011), the most commonly sung tones (in any temporal position, whether the start, middle, or end of a phrase or chant) were G and A, and the least commonly sung tone (apart from B♭) was B. A possible explanation: individual tones in the context of a chant melody differ in consonance, depending on the relationship between their audible partials and the prevailing diatonic scale. The tone G is most common because all its audible partials correspond approximately to the prevailing diatonic scale; for the same reason the tone B is least common. The chromas of the first ten harmonics of G are G, G, D, G, B, D, F, G, A, B, all of which correspond to the white-key diatonic scale; by contrast, the first ten harmonics of B are B, B, F♯, B, D♯, F♯, A, B, C♯, D♯. If this theory is correct, the consonance of the tone G in a white-tone diatonic context depends in part on the Just M3 interval between its 4th and 5th harmonics. We might therefore expect that the distance between these harmonics would influence the intonation of chant, contradicting received music-theoretical wisdom that the tuning of chant is Pythagorean.

In composition, there is a large repertoire of music intended to be played in Just intonation. The underlying idea is often that beating may be minimized or maximized by the use of particular ratios, provided the music is slow, with minimal vibrato, and perhaps computer-controlled. An example is the work of American composer David Hykes (b. 1953) and his Harmonic Choir. He developed a contemplative style of music, based on overtone singing, which he calls Harmonic Chant (1).[1] Similar examples can be found at the website of the Overtone Music Network (2).[2] The ultimate reason for Just tuning in this kind of music may not be ratios themselves but overlapping harmonics. The ratios emerge when harmonic partials line up and not because the brain is processing ratios. Another factor is the unusually slow harmonic rhythm of the music, which may remind us of a violinist tuning a P5, sustaining a dyad and aligning the partials.

[1] http://www.overtone.cc/profile/David_Hykes_Harmonic_Presence_Foundation

[2] http:// www.overtone.cc

Perception of Musical Scales

THE SUBJECTIVE SIZE OF INTERVALS IN MUSICAL CONTEXTS

The differences between pitch and frequency that we have considered can be measured by presenting two sounds in succession and asking a listener which of the two was higher. Another difference involves psychophysical scaling and the perceived size of an interval between two successive tones with different frequencies. The frequency JND of harmonic complex tones in the central musical range is almost constant when expressed in cents or as a percentage of frequency (Walliser, 1969b), implying that the most suitable physical scale for measuring musical intervals is the logarithm of frequency. On this basis, and on the basis of musical intuition, the most psychologically appropriate unit for pitch should be octaves, semitones or cents, not hertz (cycles per second, Hz).

Empirical studies have shown that the subjective size of an interval between successive pure tones corresponds neither to the frequency ratio nor its logarithm (Stevens, Volkman, & Newman, 1937). The subjective magnitude of a pure tone's pitch may vary approximately as the logarithm of frequency at higher frequencies (well above 500 Hz) and approach a linear relationship with frequency at lower frequencies (well below 500 Hz) (Zwicker, 1961). In speech, Hermes and Van Gestel (1991) confirmed that we perceive intonation along a psychoacoustic scale that lies between a linear and a logarithmic scale of frequency, reflecting the frequency selectivity of the auditory system. Based on measurements of critical bandwidth (or the effective bandwidth of individual hair cells on the basilar membrane at different characteristic frequencies), Zwicker (1961) explained the "mel scale" in terms of a critical-band function or Tonheit; Moore and Glasberg (1983) preferred the term "Equivalent-Rectangular-Bandwidth-rate "or ERB-rate. The operational definition was different, but the basic idea was the same. Although such scales apply only to pure tones, they also influence the perception of complex tones. This is further evidence that a concept of interval based on ratios is fundamentally incorrect.

Stevens (1946) presented a general theory of psychological scales and measurement with implications for psychoacoustics. Stevens' concept of "scale" differed from the musical concept, which refers to a collection of tones of nominally fixed pitches, used for making music. Stevens more generally considered the problem of perceiving points on a continuous scale relative to each other. For example, is a 100-watt light globe twice as bright as a 50-watt light globe? The answer is no, even if the 100-watt globe is producing twice as much light power, because perceived brightness is not proportional to light power.

On this basis, Stevens asked whether it is possible to measure human sensation, and defined measurement in a very general way as "the assignment of numerals to objects or events according to rules" (p. 677). He then defined four kinds of scale: nominal, ordinal, interval, and ratio. Musical scales can be of all four types. They are *nominal scales* (because scale steps have names, *do re mi*), *ordinal scales* (because scale steps are in a fixed order), *interval scales* (because intervals have the same size at different places in the scale), and *ratio scales* (because two octaves are perceived as twice as big as one octave, two semitones as twice as big as one).

Dowling (1978) found that untrained listeners can easily recognize melodies or melodic motives when the interval sizes are changed—provided the change does not affect contour, defined as the sequence of ups and downs. Although changing interval sizes can impair the recognition of familiar melodies, this finding nevertheless undermines the psychological idea of an interval scale that is based on exact interval sizes, and suggests that it may be more appropriate to regard a musical scale as an ordinal scale, or at least a scale based on approximate interval sizes.

If we built musical scales according to the Pythagorean concept of ratios, the resultant constructs would be "ratio scales" in Stevens' terminology. The listener would be sensitive to pitch ratios (e.g., tone A perceived as twice as high as tone B) whether or not she was sensitive to frequency ratios. But this idea contradicts the psychological evidence. From a psychological viewpoint, listeners can only perceive the ratio of two pitches very approximately, and the ratios do not correspond to frequency ratios; a tone an octave above another tone is not perceived to have twice the pitch, and a tone a M3 above another tone is not perceived to have 25% more pitch.

CATEGORICAL PERCEPTION OF MUSICAL PITCH

Burns and Ward (1978) presented pairs of tones to musicians and asked them to name the intervals. The participants seldom made errors. The intervals in the experiment were mistuned by as much as a quartertone, but the musicians reliably and quickly named them relative to the twelve categories of 12-EDO. For example, they quickly identified an interval of 350 cents as either m3 or M3, and an interval of 450 cents as either M3 or P4. The authors concluded that musical pitch perception is categorical. Each perceptual category has a label corresponding to a chromatic scale step and a width of about one semitone.

The perception of harmonic pitch patterns within individual complex tones is also categorical. Using Terhardt's terminology, a spectral pitch (audible partial) must correspond approximately to a harmonic of a fundamental before the auditory system perceives that partial as belonging to a complex tone with that fundamental, and the spectral pitch contributes to the salience of the virtual pitch at the fundamental. The expressions "tuning tolerance" and "category width" are quantitatively and functionally comparable.

Intonation may be deemed acceptable if listeners can correctly categorize the tones in the chromatic scale; that is, if pitches lie within category boundaries. Under good conditions for performing frequencies or perceiving pitches exactly, the categories become narrower. One may, for example, recognize an interval as a M3 if it lies between 350 and 450 cents, but it may be considered "in tune" if it lies between 380 and 430 cents.

Imagine someone listening to any such music, in any style or genre. The music suddenly stops. She then hears a pure tone with randomly selected frequency in the central musical range. In general, she will hear this tone relative to the prevailing diatonic and chromatic scales, as one of the chromatic steps (the closest) or as a natural or shifted diatonic step. Unless she has absolute pitch, or good relative pitch plus knowledge of the key in which the music was playing, she will not say "aha, that was a D♭" or "aha, that was the 4th scale degree in A♭ major," but the empirical literature on categorical perception of musical pitch suggests that she will hear the tone *as if* it were one of those things. If the tone had been tuned differently, but was still within a quartertone of D♭, she would hear it as that and may not even be aware of the mistuning. She would assume that the musicians had intended to play D♭, even if the tone was perceptibly mistuned.

This thought experiment explains why pitch is perceived categorically. In most music, mistunings of a quartertone or even a semitone are commonplace. The ear must tolerate mistunings of this order, otherwise musical appreciation would be impossible. Mistunings of partials within harmonic complex tones can be comparable; a familiar example is the stretched harmonic series in every piano tone (Martin & Ward, 1961). That can explain why partials within a harmonic complex tone can be mistuned by a quartertone or even a semitone without being separately noticed (Moore et al, 1986). Like pitches in a melody, partials are perceived categorically: either they belong to a given harmonic complex tone or they do not. This form of categorical perception was assumed by Terhardt et al. (1982) and Parncutt (1989) when considering the perception of harmonic patterns of spectral pitches, as part of a model of the pitch of complex tones that can explain aspects of the perception of harmony and tonality in music.

The empirical literature on the categorical perception of musical pitch implies that intervals do not sound as far out of tune as they really are. We are remarkably good at ignoring mistuning in performance and guessing the pitch a performer is aiming for, as if our perception were shifted toward the intended pitch of each tone (*zurecht hören*: Bruhn, 1994; Fricke, 1988; Kurth, 1931). It cannot be otherwise if the results of studies such as Devaney et al. (2012), Kopiez (2003), and Rakowski (1990) are correct. Categorical perception is like a filter that assigns out-of-tune intervals to familiar categories, as if they were in tune. This is a saving grace for amateur choirs and orchestras, if not for the best professionals.

The cited empirical literature on intonation implies that even professional musicians with the best aural abilities may have difficulty distinguishing between musical pitches and their categorical perception in musical contexts. Empirical research in the tradition of Krumhansl (1990) additionally implies that precise tuning is easier within a clear hierarchical structures, in which some pitches are more stable than others and therefore act as reference tones, which also makes intervals easier to recognize (Rakowski, 1990). This claim agrees with the personal experience of the second author while composing and rehearsing music in 19-EDO in an example of what Tymoczko (2010) calls "harmony and counterpoint in the extended common practice."

Categorical perception of musical pitch contradicts Pythagorean theory, in which musical intervals are held to have single, exact sizes. A musical interval label corresponds to a continuous range of interval sizes; any interval in the range 350 to 450 cents may be perceived as a M3. Plus or minus a quartertone might seem like a wide range of uncertainty, but that is essentially what the experiments of Burns and Ward (1978) established (see also Halpern & Zatorre, 1979; Siegel & Siegel, 1977).

CATEGORICAL PERCEPTION OF OTHER MUSICAL PARAMETERS

Intonation is one of several examples of categorical perception in music. If there is a general tendency to categorize music-structural parameters, that is another argument against the idea of intervals as ratios and for the idea of intervals as categories.

A well-known example of music-structural categorization is consonance and dissonance. There is a long tradition in music theory of grouping intervals into categories with labels such as perfect consonance, imperfect consonance, and dissonance. Chords are similarly classified as consonant or dissonant, either absolutely or relative to each other. Membership of these categories has changed from one historical period to another. For example, for much of the history of classical music, consonances were considered to contain no more than three pitch classes, whereas in the tonal music of the 20th century (blues, jazz), chords of four pitch classes (major-minor or dominant seventh, minor seventh, added sixth, major seventh) were effectively treated as consonances.

Music theorists of the Middle Ages disagreed about whether the harmonic P4 was a consonance or a dissonance. Pythagorean theory, which was originally limited to the first four numbers, regarded the 4:3 ratio as a consonance. Medieval polyphonic practice suggested otherwise: a P4 interval demanded resolution onto a more consonant third interval (imperfect consonance)—another argument against ratio theory. Today, we can assert that consonance and dissonance do not depend on ratios, because ratio theory has repeatedly failed to account for variations in consonance and dissonance. There is no clear evidence for two different M3s (5:4 and 81:64) of differing consonance in music perception or performance, and a diminished triad (5:6:7) is not more consonant than a minor triad (10:12:15).

Consonance and dissonance (like many psychological measures) are multifactorial. They depend in part on roughness (Helmholtz, 1877/1954) and harmonicity (fusion) (Stumpf, 1883), both of which depend indirectly on ratios—at least in the ideal case of simultaneities of harmonic complex tones. Consonance and dissonance also depend strongly on familiarity (Cazden, 1945), which is independent of ratios. The relative importance of the factors depends on historical and cultural context, as well as individual psychological differences, which can explain why the boundary between consonance and dissonance has changed over time (Parncutt & Hair, 2011; Tenney, 1988). At the local level in a musical passage, consonances and dissonances may still be perceived as two contrasting categories. Prefix dissonances resolve to consonances and suffix dissonances follow them—analogous to suffix neighbor-tones that sometimes follow principal melodic tones.

Another example of musical categories is rhythm. In a waltz, most sound events that we hear are categorized as falling on beats 1, 2, or 3 of the bar. Events that occur between the beats may be perceived as either beat subdivisions (eighth notes in 3/4 time) or timing deviations (anticipating and lengthening the second beat in a Viennese waltz). Psychologically, it is possible to demonstrate a category boundary between, say, 6/8 and 4/4 versions of the same melody (Schulze, 1989).

The existence of perceptual categories in the perception of rhythm and meter (Clarke, 1987) and systematic deviations from notated temporal ratios in musical rhythms (Gabrielsson, 1985) suggests that time intervals and pitch intervals have something in common; namely, they are perceived as categories rather than ratios. Analysis of timing in different rhythmic styles (Polak & London, 2014) suggests in addition that rhythms are not necessarily based on an isochronous beat. In general, dimensions of musical structures in pitch and time are learned and flexible, not physically predetermined.

Conclusion

THE NATURE OF MUSICAL INTERVALS

Musical intervals are not ratios, nor are they magical mathematical entities. They are learned, approximate, perceptual distances. They emerge from a multigenerational perceptual-historical process, mediated by the physical properties of musical tones and the physiological and psychological properties and limitations of the human auditory system.

Musical intervals belong to the subjective world of human culture, not the objective world of mathematics and physics. They are perceived between musical pitches as they occur in musical contexts, and are experienced by performers and listeners. Both pitches and the intervals between them are experiential phenomena. They are also quantitative, although they can also be described qualitatively. From a philosophical viewpoint, our detailed consideration of this problem could be regarded as a form of *quantitative phenomenology.*

These insights have profound implications for many aspects of music and musical scholarship, including practice (composition, performance), theory (music theory and analysis), sciences (acoustics, psychology, cognition, computer science, neuroscience), and humanities (history, ethnology, philosophy, theoretical sociology, aesthetics). For composers, our findings imply that ratio theory cannot be considered a scientific foundation for their art; if the music of ratio theorizing microtonalists was successful in the past, as it often was (Gilmore, 1995), other factors must explain that success. Performers who aspire to Just intonation might instead consider aspiring to (stretched) 12-EDO, given the lack of evidence that any well-known tuning system is inherently superior to any other, or that musicians of the past (singers and performers of instruments with real-time continuous pitch adjustment) deviated from 12-EDO in consistent ways. Music theorists and analysts should avoid theories based on ratios, which are misleading at best; those who document and interpret the long and fascinating history of ratio theory should avoid being seduced by it. Acousticians and psychologists should clearly separate physical effects from psychological ones, recognizing that while it may be meaningful to speak of ratios when discussing the vibrations of strings and air columns, the psychological correlate of those ratios may be something else entirely. Computer scientists should resist the temptation to develop and implement models of musical structure based on frequency ratios. Neuroscientists should keep in mind that neural synchronies, phase-locking, and resultant ratios do not necessary have correlates in conscious experience. While humanities scholars may immediately recognize the dangers of scientism and the application of quantitative theory to explain qualitative phenomena, they should be careful not to fall into the trap of taking ratio theorists at their word, just because of their enormous influence on the history of ideas.

Our approach has a modernist flavor, suggesting that all kinds of intervals, scales, and harmonies are musically possible. At first sight, that seems inconsistent with the continuing popularity of the 12-tone chromatic scale, its diatonic subsets, and the still hegemonic major-minor tonal system. But the apparently stubborn conservatism of most musical tastes can be explained another way. From a cognitive viewpoint, music processing involves a high information load that is reduced by chunking and hierarchical organization (Baddeley, 1994; Krumhansl, 1990; Tillmann et al., 2000). The foundations for the hierarchical cognitive organization of musical pitch structure are laid during early learning and enculturation (Hannon & Trainor, 2007; North, Hargreaves, & O'Neill, 2000). Such effects can explain the stability of diatonicity and major-minor tonality without any recourse to ratio theory.

The concept of interval that we have presented contradicts millennia of music theory, but it agrees with the ancient approach of Aristoxenus. Placed in a modern context, that approach has the potential to transform the foundations of music theory. Musical scales become collections of pitch categories from which different pitch combinations can be selected. Pitch-class set theory (Forte, 1977) becomes equally appropriate for tonal and atonal music, or anything in between, provided musical pitch patterns (as perceived) are defined as collections of pitch categories, rather than collections of exact pitches or (worse) frequencies. The vocabulary of melodic and voice-leading patterns, chords, and chord progressions in a given style depends on how combinations of tones corresponding to those pitch categories are perceived, not on their frequencies or notation.

The idea of intervals as ratios is still useful for answering certain questions. We may, for example, ask why Western music is based on a division of the octave into 12 equal parts, and not some other number. The reason is that ratios of small numbers match a 12-fold subdivision better than other subdivisions, while at the same time optimizing the number of tones per octave, given limitations on human capacities for information processing and additional issues of pitch discrimination and frequency adjustment in complex musical contexts. Equal subdivisions of the octave that come relatively close to small-number ratios include 5, 7, 12, 19, 22, 31, 34, 41, 53 and 72 (Gamer, 1967; Mandelbaum, 1961), of which 19, 31, and 53 are promising alternatives to 12 for modernist composition. Equal subdivisions of 5 and 7 tones per octave are already approached in some traditions; for example, Middle Eastern musics with (approximately) quartertone scales, in which the m3 interval is divided into roughly equal halves (neutral seconds), or musics with neutral third intervals that lie midway between a M3 and m3 (Spector, 1970). Another example is the Indonesian sléndro scale that divides an octave into 5 roughly equal intervals (Carterette & Kendall, 1994; Hajdu, 1993). The larger the number of tones, the greater the cognitive load, and the fewer JNDs in intervals between adjacent scale-steps. The success of such mathematical derivations does not mean the intervals themselves are ratios in an ontological sense; the derivations succeed because musical intervals are often approximations to ratios, as are intervals between partials in harmonic complex tones or harmonic complex tones in music whose partials almost coincide.

A psychologically and culturally founded approach to the theory of tonal music as it "really is" according to positivist modern science (as opposed to how it was conceived and verbalized by theorists of the past) might begin with a theory of consonance and dissonance based on psychological parameters such as roughness, harmonicity, and familiarity. It would then proceed to derive the chromatic and diatonic scales as promising solutions (but not the only musically useful ones) to the problem of how best to subdivide pitch-space under certain psychologically and culturally motivated assumptions. The theory would build up the vocabulary of tonal music by considering how tones are perceived simultaneously and sequentially in musical contexts, and comparing the predictions of the theory with the historical development of tonal-harmonic syntax in the Middle Ages, Renaissance, and Baroque periods.

THE INTRANSIGENCE OF PYTHAGOREAN BELIEFS

There are still Pythagoreans in our midst who consider intervals to be ratios, in spite of the overwhelming empirical evidence to the contrary—for example, evidence that intonation in music performance varies over a range that easily includes different theoretical ratios and may deviate systematically from them. Why do they still believe in ratios? Why is this ancient idea so resilient?

Almost a millennium after Pythagoras, Pythagorean ideas were promoted and spread in the writings of Boethius (c. 480-524), who was both a founding father of Western music theory and a major neo-Platonic philosopher (Chadwick, 1981). One might have expected the scientific revolution (15th–18th centuries) to seal the Pythagorean coffin, but if anything it did the opposite. The Pythagorean concept of intervals as ratios was championed by some of the most important figures in the history of science, including Copernicus (15th century), Galileo (16th), Descartes, Leibnitz, Newton (17th), and Euler (18th). Even Helmholtz (1877/1954) devoted long passages to ratio theory.

Why did these great scientists fail to recognize fundamental problems with ratio theory as pointed out repeatedly since Aristoxenus? The answer to this intriguing question involves histories of different kinds—political, social, cultural—as well as the history of ideas. To answer it plausibly, we need to adopt a humanities mindset and engage with non-scientific epistemological frameworks.

It is tempting to claim that scientism—the excessive emphasis on scientific approaches at the expense of others, and the belief in the inherent supremacy of the physical world (by comparison to subjectivity) and quantitative explanations (by comparison to qualitative) (Mahoney, 1989; Regelski, 1996, Sorell, 2013)—was the ultimate cause of the misleading belief in musical intervals as ratios. But in recent decades the problem has also involved a failure to acknowledge scientific methods and empirical findings. Empirical research on intonation in music performance has consistently failed to support ratio theory, even when the researchers initially set out to confirm it. To understand the historic resilience of Pythagorean thought and move on, it may help to promote a better balance between sciences and humanities (Parncutt, 2007; Snow, 1959).

The intransigence of Pythagorean thought may be founded on music's intrinsic beauty. We tend to idealize music because of the wonderful emotions it evokes. We associate positive emotions with tones that sound in tune, and hence harmonious with other tones. Research on the positive emotions evoked by sad music (Huron, 2011; Kawakami, Furukawa, Katahira, & Okanoya, 2013; Parncutt, 2014; Schubert, 1996; Vuoskoski, Thompson, McIlwain, & Eerola, 2012) suggests that the positive emotional effect of in-tuneness includes minor and dissonant harmonies. A beautiful phenomenon begs for a beautiful theory, and the elegance of Pythagorean mathematics feels intuitively right.

We can also take a broader view and consider general tendencies in the history of ideas. Music is a form of culture that shares features with the utopian Platonic and neo-Platonic thought of the ancient and medieval worlds. Pythagoreanism may be an example of the *nirvana fallacy* or *perfect solution fallacy* (Coyne, 2006): the putting of unrealistic, idealized alternatives in the place of achievable real things, or being driven by the assumption that there is a perfect solution to a particular problem. From a psychological perspective, human reasoning is based on beliefs, and simpler beliefs may be preferred because they enable more rapid or efficient responses to new situations representing either danger or opportunity. As a result, there are strong links between beliefs, desires, and intentions (Bratman, 1999; Wellman & Woolley, 1990).

It may seem far-fetched to compare Pythagorean thinking with politics. But historical political parallels may be evidence for an underlying psychosocial tendency to naively believe in simplistic theories. Consider this political example:

> Attempts to reconstruct weak and failed countries suffer from a nirvana fallacy. Where central governments are absent or dysfunctional, it is assumed that reconstruction efforts by foreign governments generate a preferable outcome. This assumption overlooks (1) the possibility that foreign government interventions can fail, (2) the possibility that reconstruction efforts can do more harm than good, and (3) the possibility that indigenous governance mechanisms may evolve that are more effective than those imposed by military occupiers." (Coyne, 2006, p. 343).

Political examples of the nirvana fallacy include the October revolution in 1917, which was motivated by the belief that communism would solve the problems of capitalism. In the 1990s, many believed that neoliberalism and "trickle-down economics" (Aghion & Bolton, 1997) would improve quality of life for all. In 2010, the theory that democracy is universally desirable led many to believe that the Arab Spring would improve living standards in Arab countries. These theories are similar to Pythagoreanism in their simple directness, and the problems that arise when they are applied in practice. The proposed solution is too simple given the complexity of the situation.

Utopian concepts are no less common in the theory and practice of music. In the 20th century, modernist utopian musical ideas grew from the ascendancy of producer autonomy over listeners' perception of musical structure (Taruskin 2009). Schoenberg famously proposed to Josef Rufer, concerning his discovery of the technique of "composing with twelve tones related only to one another," that "I have made a discovery which will ensure the supremacy of German music for the next hundred years" (Stuckenschmidt 2011, p. 277). In the optimistic post-war years following the demise of Nazism, which was itself an extreme example of a utopian worldview, Pierre Boulez criticized Schoenberg's radical musical language for not being radical enough; for him, it retained too many rhythmic and formal concepts from the classical tradition, the very features that enable most listeners to hear some coherence. These are examples of a post-Enlightenment form of utopianism that Meyer (1967) described as imagining a Golden Age, not in the remote past (as in the *Weltanshauung* of the philosophers of the ancient world or the Middle Ages), but as something still to come, to be brought about by visionary thinking in the present. They signify a refusal to let how things are stand in the way of how things ought to be.

Another reason why many still cling to Pythagorean beliefs may be their apparent success at explaining aspects of music for very long historical periods. In the history of science and ideas, there was a general reluctance to give up previously successful theories. Like Kuhn's (1962) paradigms in the history of science, Pythagoreanism may only be overthrown when its failures become intolerable and the successes of an alternative become compelling, just as classical mechanics was overthrown by quantum mechanics and relativity. As an approximation, however, the Pythagorean approach will always be useful, just as classical mechanics is still useful. Musical practice is constantly changing in accordance with changes in the economic, sociological, philosophical, historical, political, ideological, and other circumstances; humanities disciplines are consequently "devoted less to the 'advancement of knowledge' than to the propagation of moral and intellectual values" (Scruton 2007, p. 3).

Finally, we may exaggerate the "exactness" of the so-called "exact sciences." Physics is often thought of as an "exact science" because it is dominated by mathematical theory. But mathematics is also an excellent tool for dealing with *inexactness*. One approach is the order-of-magnitude estimate: the size of a quark is less than 10^{-18} m; the weak nuclear force operates over a range of 10^{-14} m; the temperature of a supernova expansion is 10^{10} Kelvin; there are 10^{14} cells in the human body; and there are 10^{80} fundamental particles in the universe. Music theory is often considered mathematical and therefore exact. In fact, the quantities considered in applied mathematics vary along a spectrum from very exact to very approximate. The number ratios that correspond to musical intervals also lie somewhere along that spectrum.

Author Note

Correspondence concerning this article should be addressed to Richard Parncutt, Centre for Systematic Musicology, University of Graz, Merangasse 70, 8010 Graz, Austria. E-mail: parncutt@uni-graz.at

References

Aghion, P., & Bolton, P. (1997). A theory of trickle-down growth and development. *Review of Economic Studies, 64*(2), 151–172.

Allanson, J. T., & Schenkel, K. D. (1965). The effect of band-limited noise on the pitch of pure tones. *Journal of Sound and Vibration, 2*(4), 402–408.

Ambrazevičius, R., & Wiśniewska, I. (2009). Tonal hierarchies in Sutartinės. *Journal of Interdisciplinary Music Studies, 3*(1–2), 45–55.

Anderson, B. E., & Strong, W. J. (2005). The effect of inharmonic partials on pitch of piano tones. *Journal of the Acoustical Society of America, 117*(5), 3268–3272.

Araya-Salas, M. (2012). Is birdsong music? Evaluating harmonic intervals in songs of a Neotropical songbird. *Animal Behaviour, 84*(2), 309–313.

Arnold, W. J. (1985). L'intonation juste dans la théorie ancienne de l'Inde: Ses applications aux musiques modale et harmonique [Just intonation in the history of Indian theory: Its applications to modal and harmonic musics]. *Revue de Musicologie, 71,* 11–38.

Arom, S., Léothard, G., & Voisin, F. (1997). Experimental ethnomusicology: An interactive approach to the study of musical scales. In I. Deliège & J. Sloboda (Eds.), *Perception and cognition of music* (pp. 3–30). Hove, East Sussex, UK: Psychology Press.

Aures, W. (1985). Procedure for calculating the sensory euphony of arbitrary sound signals. *Acustica, 59*(2), 130–141.

Ayers, T., Aeschbach, S., & Walker, E. L. (1980). Psychoacoustic and experiential determinants of tonal consonance. *Journal of Auditory Research, 20*(1), 31–42.

Baddeley, A. (1994). The magical number seven: Still magic after all these years? *Psychological Review, 101*(2), 353–356.

Barbour, J. M. (1938). Just intonation confuted. *Music and Letters, 19,* 48–60.

Barbour, J. M. (2004). *Tuning and temperament: A historical survey.* Mineola, NY: Dover.

Barlow, C. (1987). Two essays on theory. *Computer Music Journal, 11,* 44–59.

Békésy, G. von (1963). Three experiments concerned with pitch perception. *Journal of the Acoustical Society of America, 35*(4), 602–606.

Bibas, A. B., Xenellis, J., Michaels, L., Anagnostopoulou, S., Ferekidis, E., & Wright, A. (2008). Temporal bone study of development of the organ of Corti: Correlation between auditory function and anatomical structure. *Journal of Laryngology and Otology*, *122*, 336–342.

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.

Bigand, E., Parncutt, R., & Lerdahl, F. (1996). Perception of musical tension in short chord sequences: Influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception and Psychophysics, 58*, 125–141.

Bower, C. M. (2002). The transmission of ancient music theory into the Middle Ages. In Christensen (Ed.), *Cambridge history of Western music theory* (pp. 136–167). Cambridge, UK: Cambridge University Press.

Bratman, M. E. (1999). Intention, plans, and practical reason. Cambridge, MA: Harvard University Press.

Bregman, A. (1990). *Auditory scene analysis.* Cambridge, MA: MIT Press.

Brown, J. C., & Vaughn, K. V. (1994). Pitch center of stringed instrument vibrato tones. *Journal of the Acoustical Society of America, 100*, 1728–1735.

Bruhn, H. (1994). *Wahrnehmung von Musik. Eine Allgemeine Musiklehre aus der Sicht von Psychologie und Musikgeschichte* [Perception of music: A general music theory from the perspective of psychology and music history]. Kiel: Christian-Albrechts-Universität.

Burns, E. (1999). Intervals, scales, and tuning. In D. Deutsch (Ed.), *Psychology of music* (2nd ed., pp. 215–264). San Diego, CA: Elsevier.

Burns, E. M. (1982). Pure-tone pitch anomalies. I. Pitch-intensity effects and diplacusis in normal ears. *Journal of the Acoustical Society of America, 72*(5), 1394–1402.

Burns, E. M., & Ward, W. D. (1978). Categorical perception—Phenomenon or epiphenonon: Evidence from experiments in the perception of melodic musical intervals. *Journal of the Acoustic Society of America, 63,* 456–468.

Burns, E. M., & Ward, W. D. (1982). Intervals, scales, and tuning. In D. Deutsch (Ed.), *Psychology of music* (pp. 241–269). San Diego: Elsevier.

Caleon, I., & Ramanathan, S. (2008). From music to physics: The undervalued legacy of Pythagoras. *Science and Education, 17*(4), 449–456.

Cariani, P. (1999). Temporal coding of periodicity pitch in the auditory system: An overview. *Neural Plasticity, 6*(4), 147–172.

Carlyon, R. P. (1991). Discriminating between coherent and incoherent frequency modulation of complex tones. *Journal of the Acoustical Society of America, 89*(1), 329–340.

Carterette, E. C., & Kendall, R. A. (1994). On the tuning and stretched octave of Javanese gamelans. *Leonardo Music Journal, 4,* 59–68.

Cazden, N. (1945). Musical consonance and dissonance: A cultural criterion. *Journal of Aesthetics and Art Criticism, 4*(1), 3–11.

Chadwick, H. (1981). *Boethius: The consolations of music, logic, theology, and philosophy.* Oxford, UK: Oxford University Press.

Chalmers, J. H. (1993). *Divisions of the tetrachord: A prolegomenon to the construction of musical scales.* Hanover, NH: Frog Peak Music.

Cheung, K. K. (2013). A computational approach towards just intonation for Renaissance music. *World Journal of Computer Application and Technology, 1*(1), 1–5.

Cheveigné, A. de (2003). Time-domain auditory processing of speech. *Journal of Phonetics, 31*(3), 547–561.

Clarke, E. F. (1987). Categorical rhythm perception. An ecological perspective. In A. Gabrielsson (Ed.), *Action and perception in rhythm and music* (pp. 19-33). Stockholm: Royal Swedish Academy of Music.

Cohen, A. J., Thorpe, L. A., & Trehub, S. E. (1987). Infants' perception of musical relations in short transposed tone sequences. *Canadian Journal of Psychology, 4*(1), 33-47.

Collins, T., Tillmann, B., Barrett, F. S., Delbe, C., & Janata, P. (2014). A combined model of sensory and cognitive representations underlying tonal expectations in music: From audio signals to behavior. *Psychological Review, 121*(1), 33-65.

Corso, J. F. (1954). Scale position and performed melodic octaves. *Journal of Psychology, 37,* 297–305.

Coyne, C. J. (2006). Reconstructing weak and failed states: Foreign intervention and the nirvana fallacy. *Foreign Policy Analysis 2*(4), 343-360.

Cynx, J., & Shapiro, M. (1986). Perception of the missing fundamental by a species of songbird (Sturnus vulgaris). *Journal of Comparative Psychology, 100,* 356–360.

Dalby, B. F. (1992). A computer-based training program for developing harmonic intonation discrimination skill. *Journal of Research in Music Education, 40*(2), 139-152.

DeCasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development, 9*(2), 133-150.

Delviniotis, D., Kouroupetroglou, G., & Theodoridis, S. (2008). Acoustic analysis of musical intervals in modern Byzantine Chant scales. *Journal of the Acoustical Society of America, 124*(4), EL262-EL269.

Devaney, J., Mandel, M. I., Ellis, D. P. W., & Fujinaga, I. (2011). Automatically extracting performance data from recordings of trained singers. *Psychomusicology, 21*(1–2), 108–136.

Devaney, J., Mandel, M. I., & Fujinaga, I. (2012). A study of intonation in three-part singing using the Automatic Music Performance Analysis and Comparison Toolkit (AMPACT). In F. Gouyon et al., (Ed.), *Proceedings of the 13th International Society for Music Information Retrieval Conference* (pp. 511–516). Porto, Portugal: FEUP.

Dobbins, P. A., & Cuddy, L. L. (1982). Octave discrimination: An experimental confirmation of the "stretched"subjective octave. *Journal of the Acoustical Society of America, 72*(2), 411–415.

Don, G. W. (2001). Brilliant colors provocatively mixed: Overtone structures in the music of Debussy. *Music Theory Spectrum, 23*(1), 61–73.

Doolittle, E., & Brumm, H. (2012). O Canto do Uirapuru: Consonant intervals and patterns in the song of the musician wren. *Journal of Interdisciplinary Music Studies, 6*(1), 55–85.

Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review, 85*(4), 341–354.

Duffin, R. W. (2007). *How equal temperament ruined harmony (and why you should care).* New York: Norton.

Duke, R. A. (1985). Wind instrumentalists' intonational performance of selected musical intervals. *Journal of Research in Music Education, 33*(2), 101–111.

Egan, J. P., & Meyer, D. R. (1950). Changes in pitch of tones of low frequency as a function of the pattern of excitation produced by a band of noise. *Journal of the Acoustical Society of America, 22*(6), 827–833.

Erlich, Paul (1998). Tuning, tonality, and twenty-two-tone temperament. *Xenharmonikôn, 17,* 12–40.

Fauvel, J., Flood, R., & Wilson, R. J. (2003). *Music and mathematics: From Pythagoras to fractals*. Oxford: Oxford University Press

Fay, R. (1984). Temporal processing by the auditory system of fishes. In A. Michelsen (Ed.), *Time resolution in auditory systems* (pp. 28–57). Berlin: Springer-Verlag.

Fokker, A. D. (1949). *Just intonation and the combination of harmonic diatonic melodic groups.* The Hague, Netherlands: Martinus Nijhoff.

Fonseca, P. J., Münch, D., & Hennig, R. M. (2000). Auditory perception: How cicadas interpret acoustic signals. *Nature, 405*(6784), 297–298.

Fonville, J. (1991). Ben Johnston's extended just intonation: A guide for interpreters. *Perspectives of New Music, 29,* 106–137.

Forte, A. (1977). *The structure of atonal music.* New Haven, CT: Yale University Press.

Foss, A. H., Altschuler, E. L., & James, K. H. (2007). Neural correlates of the Pythagorean ratio rules. *Neuroreport, 18*(15), 1521–1525.

Fricke, J. P. (1988). Klangbreite und Tonempfindung. Bedingungen kategorialer Wahrnehmung aufgrund experimenteller Untersuchung der Intonation [Sound width and tone sensation: Conditions for categorical perception based on experimental investigation of intonation]. *Jahrbuch Musikpsychologie, 5*, 67–87.

Frosch, R. (1993). *Mitteltönig ist schöner*! *Studien über Stimmungen von Musikinstrumenten* [Meantone is nicer! Studies on tunings of musical instruments] (2nd ed.). Bern, Switzerland: Lang.

Fyk, J. (1995). *Melodic intonation, psychoacoustics, and the violin*. Zielona Góra, Poland: Organon.

Gabrielsson, A. (1985). Interplay between analysis and synthesis in studies of music performance and music experience. *Music Perception, 3*, 59–86.

Gamer, C. (1967). Some combinational resources of equal-tempered systems. *Journal of Music Theory, 11*(1), 32–59.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston, MA: Houghton Mifflin.

Gill, K. Z., & Purves, D. (2009). A biological rationale for musical scales. *PLoS One, 4*(12), e8144.

Gilmore, B. (1995). Changing the metaphor: Ratio models of musical pitch in the work of Harry Partch, Ben Johnston, and James Tenney. *Perspectives of New Music, 33,* 458–503.

Godwin, J. (1992). *The harmony of the spheres: The Pythagorean tradition in music.* Rochester, VT: Inner Traditions.

Goldacre, B. (2010). *Bad science: Quacks, hacks, and big pharma flacks.* London, UK: McClelland & Stewart.

Gratzki, B. (1993). *Die reine Intonation im Chorgesang* (Vol. 70). Bonn, Germany: Verlag für. Systematische Musikwissenschaft.

Greene, P. C. (1937). Violin intonation. *Journal of the Acoustical Society of America, 8*(3), 207–207.

Hagerman, B., & Sundberg, J. (1980). Fundamental frequency adjustment in barbershop singing. *Journal of Research in Singing, 4*, 3–17.

Hajdu, G. (1993). Low energy and equal spacing: The multifactorial evolution of tuning systems. *Journal of New Music Research, 22*(4), 319–333.

Hall, D. E. (1974). Quantitative evaluation of musical scale tunings. *American Journal of Physics, 42*(7), 543–552.

Halpern, A. R., & Zatorre, R. J. (1979). Identification, discrimination, and selective adaptation of simultaneous musical intervals. *Journal of the Acoustical Society of America, 65*(S1), S40–S40.

Hannon, E. E., Soley, G., & Levine, R. S. (2011). Constraints on infants' musical rhythm perception: Effects of interval ratio complexity and enculturation. *Developmental Science, 14*(4), 865–872.

Hannon, E. E., & Trainor, L. J. (2007). Music acquisition: Effects of enculturation and formal training on development. *Trends in Cognitive Sciences, 11*(11), 466–472.

Hasegawa, R. (2006). Tone representation and just intervals in contemporary music. *Contemporary Music Review, 25*(3), 263–281.

Havrøy, F. (2013). 'You cannot just say: "I am singing the right note"': Discussing intonation issues with Neue Vocalsolisten Stuttgart. *Music and Practice, 1*(1). Retrieved from musicandpractice.org.

He, C., & Trainor, L. J. (2009). Finding the pitch of the missing fundamental in infants. *Journal of Neuroscience, 29*(24), 7718–8822.

Heffner, H., & Whitfield, I. C. (1976). Perception of the missing fundamental by cats. *Journal of the Acoustical Society of America, 59,* 915–919.

Heinbach, W. (1988). Aurally adequate signal representation: The part-tone-time-pattern. *Acustica, 67*(2), 113–121.

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

Heman, C. (1964). *Intonation auf Streichinstrumenten: Melodisches und harmonisches Hören* [Intonation of string instruments: Melodic and harmonic hearing]. Basel: Bärenreiter-Verlag.

Hepper, P. G., & Shahidullah, B.S. (1994). Development of fetal hearing. *Archives of Disease in Childhood, 71,* F81–F87.

Hermes, D. J., & Van Gestel, J. C. (1991). The frequency scale of speech intonation. *Journal of the Acoustical Society of America, 90*(1), 97–102.

Hesse, A. (1987). Zur Tonhöhenverschiebung von Sinustönen durch Störgeräusche [Pitch shift of pure tones due to masking noise]. *Acustica, 62*(4), 264–281.

Honingh, A., & Bod, R. (2005). Convexity and the well-formedness of musical objects. *Journal of New Music Research, 34*(3), 293–303.

Houtsma, A. J. M., Rossing, T. D., & Wagenaars, W. M. (1987). *Auditory demonstrations* [CD with explanatory booklet]. New York: Acoustical Society of America.

Huron, D. (1994). Interval-class content in equally tempered pitch-class sets: Common scales exhibit optimum tonal consonance. *Music Perception, 11*, 289–305.

Huron, D. (2001). Tone and voice: A derivation of the rules of voice-leading from perceptual principles. *Music Perception*, *19*, 1-64.

Huron, D. (2011). Why is sad music pleasurable? A possible role for prolactin. *Musicae Scientiae, 15*(2), 146–158.

Hutchinson, W., & Knopoff, L. (1978). The acoustic component of Western consonance. *Journal of New Music Research, 7*(1), 1–29.

James, J. (1993). *The music of the spheres: Music, science, and the natural order of the universe.* New York: Grove.

Kantorski, V. J. (1986). String instrument intonation in upper and lower registers: The effects of accompaniment. *Journal of Research in Music Education, 34*(3), 200–210.

Karrick, B. (1998). An examination of the intonation tendencies of wind instrumentalists based on their performance of selected harmonic musical intervals. *Journal of Research in Music Education, 46*(1), 112–127.

Kawakami, A., Furukawa, K., Katahira, K., & Okanoya, K. (2013). Sad music induces pleasant emotion. *Frontiers in Psychology, 4*, 311.

Kirck, G. T. (1987). Computer realization of extended just intonation compositions. *Computer Music Journal, 11*(1), 69–75.

Koelsch, S., & Siebel, W. A. (2005). Towards a neural basis of music perception. *Trends in Cognitive Sciences, 9*(12), 578–584.

Kopiez, R. (2003). Intonation of harmonic intervals: Adaptability of expert musicians to equal temperament and just intonation. *Music Perception, 20*, 383–410.

Krahenbuehl, D., & Schmidt, C. (1962). On the development of musical systems. *Journal of Music Theory, 6,* 32–65.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.

Kuhn, T. S. (1962). *The structure of scientific revolutions.* Chicago, IL: University of Chicago Press.

Kurth, E. (1931). *Musikpsychologie*. Berlin: Hesse.

Laden, B., & Keefe, D. H. (1989). The representation of pitch in a neural net model of chord classification. *Computer Music Journal, 13*(4), 12–26.

Langner, G. (1997). Neural processing and representation of periodicity pitch. *Acta Oto-Laryngologica*, *117*(sup532), 68–76.

Leissinger, U. (1994). *Leibniz-Reflexe in der deutschen Musiktheorie des 18. Jahrhunderts* [Leibniz reflexes in the German music theory of the 18th century]. Würzburg: Königshausen und Neumann.

Licklider, J. C. R. (1957). Effects of changes in the phase pattern upon the sound of a 16-harmonic tone. *Journal of the Acoustical Society of America, 29*(6), 780–780.

Loosen, F. (1993). Intonation of solo violin performance with reference to equally tempered, Pythagorean, and just intonations. *Journal of the Acoustical Society of America, 93*(1), 525–539.

Loosen, F. (1995). The effect of musical experience on the conception of accurate tuning. *Music Perception, 12*, 291–306.

Lukács, G. (1983). *Record of a life* (I. Eörsi, Ed., R. Livingstone, Trans.). London, UK: Verso.

Madsen, C. K., & Geringer, J. M. (1981). Discrimination between tone quality and intonation in unaccompanied flute/oboe duets. *Journal of Research in Music Education, 29*(4), 305–313.

Mahoney, M. J. (1989). Scientific psychology and radical behaviorism: Important distinctions based in scientism and objectivism. *American Psychologist*, *44*(11), 1372.

Mandelbaum, J. (1961). *Multiple division of the octave and the tonal resources of 19-tone temperament* (Unpublished doctoral dissertation). Indiana University.

Mantel, G. (2005). *Intonation: Spielräume für Streiche.* Mainz: Schott.

Martin, D. W., & Ward, W. D. (1961). Subjective evaluation of musical scale temperament in pianos. *Journal of the Acoustical Society of America, 33*(5), 582–585.

Mason, J. A. (1960). Comparison of solo and ensemble performances with reference to Pythagorean, just, and equitempered intonations. *Journal of Research in Music Education, 8*, 31–38.

Mathews, M. V., Pierce, J. R., Reeves, A., & Roberts, L. A. (1988). Theoretical and experimental explorations of the Bohlen–Pierce scale. *Journal of the Acoustical Society of America, 84*(4), 1214–1222.

Mathiesen, T. J. (1999). *Apollo's lyre: Greek music and music theory in antiquity and the Middle Ages.* Lincoln, NE: University of Nebraska Press.

McAdams, S. (1989). Segregation of concurrent sounds. I: Effects of frequency modulation coherence. *Journal of the Acoustical Society of America, 86*(6), 2148–2159.

Meddis, R., & O'Mard, L. (1997). A unitary model of pitch perception. *Journal of the Acoustical Society of America*, *102*(3), 1811–1820.

Menke, D. (2009). *Musikalische Aspekte von Intonation*. Hamburg: Diplomica.

Meyer, L. B. (1967). *Music, the arts and ideas*. Chicago, IL: University of Chicago Press.

Miller, G. A., & Heise, G. A. (1950). The trill threshold. *Journal of the Acoustical Society of America, 22* (5), 637–638.

Milne, A., Sethares, W., & Plamondon, J. (2008). Tuning continua and keyboard layouts. *Journal of Mathematics and Music, 2*(1), 1–19.

Moore, B. C. J. (1989). *An introduction to the psychology of hearing* (3rd ed.). Bingley, UK: Emerald.

Moore, B. C. J. (2012). *An introduction to the psychology of hearing* (6th ed.). Bingley, UK: Emerald.

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

Moore, B. C., Glasberg, B. R., & Peters, R. W. (1986). Thresholds for hearing mistuned partials as separate tones in harmonic complexes. *Journal of the Acoustical Society of America, 80*(2), 479–483.

Morrison, S. J. (2000). Effect of melodic context, tuning behaviors, and experience on the intonation accuracy of wind players. *Journal of Research in Music Education, 48*(1), 39–51.

Morrison, S. J., & Fyk, J. (2002). Intonation. In R. Parncutt & G.E. McPherson (Eds.), *Science and psychology of music performance* (pp. 183–197). New York: Oxford University Press.

Narmour, E. (1992). The *analysis and cognition of melodic complexity: The implication-realization model.* Chicago, IL: University of Chicago Press.

Nickerson, J. F. (1949). Intonation of solo and ensemble performance of the same melody. *Journal of the Acoustical Society of America, 21*(6), 593–595.

North, A. C., Hargreaves, D. J., & O'Neill, S. A. (2000). The importance of music to adolescents. *British Journal of Educational Psychology, 70*(2), 255–272.

O'Keefe, V. (1975). Psychological preference for harmonized musical passages in the just and equally-tempered systems. *Perception and Motor Skills, 40,* 192–194.

Orlikoff, R. F., & Kahane, J. C. (1991). Influence of mean sound pressure level on jitter and shimmer measures. *Journal of Voice, 5*(2), 113–119.

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

Parncutt, R. (2001). Critical comparison of acoustical and perceptual theories of the origin of musical scales. In D. Bonsi, D. Gonzalez, & D. Stanzial (Eds.), *Proceedings of the International Symposium of Musical Acoustics.* Venice: Fondazione Scuola di San Giorgio.

Parncutt, R. (2003). Accents and expression in piano performance. In K. W. Niemöller (Ed.), *Perspektiven und Methoden einer Systemischen Musikwissenschaft* (Festschrift Fricke) (pp. 163–185). Frankfurt/Main, Germany: Peter Lang.

Parncutt, R. (2007). Systematic musicology and the history and future of western musical scholarship. *Journal of Interdisciplinary Music Studies, 1*, 1–32.

Parncutt, R. (2009). Prenatal and infant conditioning, the mother schema, and the origins of music and religion. *Musicae Scientiae*, 13 (2_suppl), 119–150.

Parncutt, R. (2012). Psychoacoustics and cognition for musicians. In A. Brown (Ed.), Sound musicianship: Understanding the crafts of music (pp. 76–87). Cambridge, UK: Cambridge Scholars Publishing.

Parncutt, R. (2014). The emotional connotations of major versus minor tonality: One or more origins? *Musicae Scientiae, 18*(3), 324–353.

Parncutt, R., & Cohen, A. J. (1995). Identification of microtonal melodies: Effects of scale-step size, serial order, and training. *Perception and Psychophysics, 57*, 835–846.

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

Parncutt, R. & Prem, D. (2008, August). *The relative prevalence of Medieval modes and the origin of the leading tone*. Poster presentation at International Conference on Music Perception and Cognition, Sapporo, Japan.

Parncutt, R., Reisinger, D., Fuchs, A., & Kaiser, F. (submitted). *Consonance and prevalence of sonorities in western polyphony: Roughness, harmonicity, familiarity, evenness, diatonicity.*

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience, 6*(7), 674–681.

Patel, A. D., & Balaban, E. (2001). Human pitch perception is reflected in the timing of stimulus-related cortical activity. *Nature Neuroscience*, *4*(8), 839–844.

Patterson, R. D. (1973). The effects of relative phase and the number of components on residue pitch. *Journal of the Acoustical Society of America, 53*(6), 1565-1572.

Patterson, R. D. (1987). A pulse ribbon model of monaural phase perception. *Journal of the Acoustical Society of America, 82*(5), 1560–1586.

Pell, M. D. (2006). Cerebral mechanisms for understanding emotional prosody in speech. *Brain and Language, 96*(2), 221–234.

Plomp, R. (1967). Beats of mistuned consonances. *Journal of the Acoustical Society of America, 42,* 462–474.

Plomp, R., & Steeneken, H. J. M. (1969). Effect of phase on the timbre of complex tones. *Journal of the Acoustical Society of America, 46*(2B), 409–421.

Polak, R., & London, J. (2014). Timing and meter in Mande drumming from Mali. *Music Theory Online, 20*(1).

Prame, E. (1997). Vibrato extent and intonation in professional Western lyric singing. *Journal of the Acoustical Society of America, 102*, 616–621.

Preisler, A., & Schmidt, S. (1998). Spontaneous classification of complex tones at high and ultrasonic frequencies in the bat, Megaderma lyra. *Journal of the Acoustical Society of America, 103*(5), 2595–2607.

Pujol, R., & Lavigne-Rebillard, M. (1995). Sensory and neural structures in the developing human cochlea. *International Journal of Pediatrics and Otorhinolaryngology, 32,* S177–182.

Rakowski, A. (1976). Tuning of isolated musical intervals. *Journal of the Acoustical Society of America, 59*, S50.

Rakowski, A. (1985). The perception of musical intervals by music students. *Bulletin of the Council for Research in Music Education, 85*, 175–186.

Rakowski, A. (1990). Intonation variants of musical intervals in isolation and in musical contexts. *Psychology of Music, 18*, 60–72.

Rakowski, A. (1994). Musician's tendency to stretch larger-than-octave melodic intervals. *Journal of the Acoustical Society of America, 96*, 3301.

Rasch, R. A. (1979). Synchronization in performed ensemble music. *Acustica, 43*(2), 121–131.

Reed, E. S. (1996). *Encountering the world: Toward an ecological psychology.* New York: Oxford University Press.

Regelski, T. A. (1996). Scientism in experimental music research. *Philosophy of Music Education Review, 4,* 3–19.

Richards, D. S., Frentzen, B., Gerhardt, K. J., McCann, M. E., & Abrams, R. M. (1992). Sound levels in the human uterus. *Obstetrics and Gynecology, 80,* 186–190.

Ritsma, R. J. (1967). Frequencies dominant in the perception of the pitch of complex sounds. *Journal of the Acoustical Society of America, 42*(1), 191–198.

Roberts, L. A., & Mathews, M. V. (1984). Intonation sensitivity for traditional and nontraditional chords. *Journal of the Acoustical Society of America, 75*(3), 952–959.

Rosner, B. S. (1999). Stretching and compression in the perception of musical intervals. *Music Perception, 17*, 101–113.

Russell, B. (1945). *A history of Western philosophy.* New York: Simon & Schuster.

Schellenberg, E. G., & Trehub, S. E. (1996). Natural musical intervals: Evidence from infant listeners. *Psychological Science, 7*(5), 272–277.

Schiemer, G., Alves, B., Taylor, S. & Havryliv, M. (2003). Pocket Gamelan: Developing the instrumentarium for an extended harmonic universe. In B. Tan. (Ed.), *Proceedings of the International Computer Music Conference* (pp. 329–332). San Francisco, CA: International Computer Music Association.

Schubert, E. (1996). Enjoyment of negative emotions in music: An associative network explanation. *Psychology of Music, 24*(1), 18–28.

Schubert, E. D. (1950). The effect of a thermal masking noise on the pitch of a pure tone. *Journal of the Acoustical Society of America, 22*(4), 497–499.

Schulze, H. H. (1989). Categorical perception of rhythmic patterns. *Psychological Research, 51*(1), 10–15.

Schwartz, D. A., Howe, C. Q., & Purves, D. (2003). The statistical structure of human speech sounds predicts musical universals. *Journal of Neuroscience, 23*(18), 7160–7168.

Scruton, R. (2007). *Philosophy: Principles and problems.* London, UK: Continuum.

Sethares, W. A. (1994). Adaptive tunings for musical scales. *Journal of the Acoustical Society of America, 96*(1), 10–18.

Shonle, J. I., & Horan, K. E. (1976). Trill threshold revisited. *Journal of the Acoustical Society of America, 59* (2), 469–471

Siegel, J. A., & Siegel, W. (1977). Categorical perception of tonal intervals: Musicians can't tell sharp from flat. *Perception and Psychophysics, 21*(5), 399–407.

Snow, C. P. (1959). *The two cultures and the scientific revolution.* New York: Cambridge University Press.

Sonntag, B. (1983). Tonhöhenverschiebung von Sinustönen durch Terzrauschen bei unterschiedlichen Frequenzlagen [Pitch shift of pure tones by third-band noise in different frequency regions]. *Acustica, 53*(4), 218–223.

Sorell, T. (2013). *Scientism: Philosophy and the infatuation with science*. Abingdon, UK: Routledge.

Spector, J. (1970). Classical 'ud music in Egypt with special reference to Maqamat. *Ethnomusicology, 14*(2), 243–257.

Stevens, S. S. (1946). On the theory of scales of measurement. *Science* (New Series), *103* (2684), 677–680.

Stevens, S. S., Volkman, J., & Newman, E. B. (1937). A scale for the measurement of the psychological magnitude pitch. *Journal of the Acoustical Society of America, 8*(3), 185–190.

Stoll, G. (1985). Pitch shift of pure and complex tones induced by masking noise. *Journal of the Acoustical Society of America, 77*(1), 188–192.

Stuckenschmidt, H. H. (2011). *Schoenberg: His life, world, and work* (H. Searle, Trans.). London, UK: Overture Publishing.

Stumpf, C. (1883). *Tonpsychologie* [Tone psychology]. Leipzig: Hirzel.

Sundberg, J. (1982). In tune or not? A study of fundamental frequency in music practice. *Department of Speech, Music and Hearing: Quarterly Progress and Status Report* (STL-QPSR), *23*(1), 49–78.

Sundberg, J. (1998). Expressivity in singing. A review of some recent investigations. *Logopedics Phonatrics Vocology, 23*(3), 121–127.

Sundberg, J. E. F., & Lindqvist, J. (1973). Musical octaves and pitch. *Journal of the Acoustical Society of America, 54*(4), 922–929.

Sundberg, J., Askenfelt, A., & Frydén, L. (1983). Musical performance: A synthesis-by-rule approach. *Computer Music Journal, 7*(1), 37–43.

Taruskin, R. (2009). *The danger of music and other anti-utopian essays*. Berkeley, CA: University of California Press.

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

Terhardt, E. (1972). Zur Tonhöhenwahrnehmung von Klängen [Pitch perception of complex sounds]. *Acustica, 26,* 173–199.

Terhardt, E. (1974). Pitch of pure tones: Its relation to intensity. In E. Zwicker & E. Terhardt (Eds.), *Facts and models in hearing* (pp. 353-360). Berlin: Springer.

Terhardt, E. (1976). Ein psychoakustisch begründetes Konzept der musikalischen Konsonanz [A psychoacoustically grounded concept of musical consonance]. *Acustica, 36*, 121–137.

Terhardt, E. (1979). Calculating virtual pitch. *Hearing Research, 1*(2), 155–182.

Terhardt, E. (1987). Gestalt principles and music perception. In W. A. Yost & C. S. Watson (Eds.), *Auditory processing of complex sounds* (pp. 157–166). Hillsdale, NJ: Erlbaum.

Terhardt, E. (1988). Psychophysikalische Grundlagen der Beurteilung musikalischer Klänge [Psychophysical foundations of musical sound judgments]. In J. Meyer (Ed.), *Qualitätsaspekte bei Musikinstrumenten* (pp. 1–15). Celle: Moeck.

Terhardt, E. (1998). *Akustische Kommunikation* [Acoustic communication]. Berlin: Springer.

Terhardt, E., & Fastl, H. (1971). Zum Einfluss von Störtönen und Störgeräuschen auf die Tonhöhe von Sinustönen [The effect of masking tones and noises on the pitch of pure tones]. *Acustica, 25*, 53–61.

TERHARDT, E., STOLL, G., & SEEWANN, M. (1982). Pitch of complex signals according to virtual-pitch theory: Tests, examples, and predictions. *Journal of the Acoustical Society of America, 71*(3), 671–678.

THOMPSON, J., HANKINSON, A., & FUJINAGA, I. (2011). Searching the Liber Usualis: Using COUCHDB and ELASTICSEARCH to query graphical music documents. In A. Klapuir & C. Leider (Eds.), *Proceedings of the 12th International Society for Music Information Retrieval Conference*. Retrieved from http://ismir2011.ismir.net/

TILLMANN, B., BHARUCHA, J. J., & BIGAND, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review, 107*(4), 885–913.

TODD, P. M., & LOY, D. G. (Eds., 1991). *Music and connectionism*. Cambridge, MA: MIT Press.

TRAINOR, L. (1997). Effects of frequency ratio on infants' and adults' discrimination of simultaneous intervals. *Journal of Experimental Psychology: Human Perception and Performance, 23*(5), 1427–1438.

TREHUB, S. E., & HANNON, E. E. (2006). Infant music perception: Domain-general or domain-specific mechanisms? *Cognition, 100*(1), 73–99.

TYMOCZKO, D. (2010). *A geometry of music: Harmony and counterpoint in the extended common practice*. Oxford, UK: Oxford University Press.

VERSCHUURE, J., & VAN MEETEREN, A. A. (1975). The effect of intensity on pitch. *Acta Acustica, 32*(1), 33–44.

VOGEL, M. (1955). *Die Zahl Sieben in der spekulativen Musiktheorie* [The number seven in speculative music theory] (Unpublished doctoral dissertation). University of Bonn, Germany.

VOGEL, M. (1975). *Die Lehre von den Tonbeziehungen* [The theory of tone relationships]. Bonn, Germany: Verlag für. Systematische Musikwissenschaft.

VOS, J. (1986). Purity ratings of tempered fifths and major thirds. *Music Perception, 3*, 221–257.

VOS, J., & VAN VIANEN, B. G. (1985). Thresholds for discrimination between pure and tempered intervals: The relevance of nearly coinciding harmonics. *Journal of the Acoustical Society of America, 77*(1), 176–187.

VOS, P. G., & TROOST, J. M. (1989). Ascending and descending melodic intervals: Statistical findings and their perceptual relevance. *Music Perception, 6,* 383–396.

VUOSKOSKI, J. K., THOMPSON, W. F., MCILWAIN, D., & EEROLA, T. (2012). Who enjoys listening to sad music and why? *Music Perception, 29*, 311–317.

VURMA, A., & ROSS, J. (2006). Production and perception of musical intervals. *Music Perception, 23*, 331–344.

WALLISER, K. (1969a). Über die Spreizung von empfundenen Intervallen gegenüber mathematisch harmonischen Intervallen bei Sinustönen [Stretching of perceived intervals relative to mathematically harmonic intervals between pure tones]. *Frequenz, 23,* 139–143.

WALLISER, K. (1969b). Zur Unterschiedsschwelle der Periodentonhöhe [The difference threshold of periodicity pitch]. *Acustica, 21*, 329–336.

WEBSTER, J. C., & SCHUBERT, E. D. (1954). Pitch shifts accompanying certain auditory threshold shifts. *Journal of the Acoustical Society of America, 26*(5), 754–758.

WEBSTER, J. C., MILLER, P. H., THOMPSON, P. O., & DAVENPORT, E. W. (1952). The masking and pitch shifts of pure tones near abrupt changes in a thermal noise spectrum. *Journal of the Acoustical Society of America, 24*(2), 147–152.

WELLMAN, H. M., & WOOLLEY, J. D. (1990). From simple desires to ordinary beliefs: The early development of everyday psychology. *Cognition, 35*(3), 245–275.

WHITCOMB, B. (2007). Improving intonation. *American String Teacher, 57*(4), 42–45.

WILLIAMSON, C. (1939). A design for a keyboard instrument in just intonation. *Journal of the Acoustical Society of America, 11*(2), 216–218.

WOLFE, J., & SCHUBERT, E. (2008). *They were playing our song: A psycho-acoustic explanation of why non-vocal musical instruments determined how we sing*. Paper presentation at the 10th International Conference on Music Perception and Cognition, Sapporo, Japan.

XU, Y., & EMILY WANG, Q. (2001). Pitch targets and their realization: Evidence from Mandarin Chinese. *Speech communication, 33*(4), 319–337.

ZATORRE, R. J., EVANS, A. C., & MEYER, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience, 14*(4), 1908–1919.

ZWICKER, E. (1961). Subdivision of the audible frequency range into critical bands (Frequenzgruppen). *Journal of the Acoustical Society of America, 33*(2), 248–248.