Commentary on Buechele et al. (2024): Communicating Across the Divide – a Place for Physics in Music?

JOHN M. MCBRIDE [1]
Center for Soft and Living Matter, Institute for Basic Science, South Korea

ABSTRACT: The theory of Jesse Berezovsky (2019) is a rare foray of a physicist into the territory of music science. In their follow-up article in *Empirical Musicology Review*, Ryan Buechele, Alex Cooke, and Jesse Berezovsky (2024) show how the evolution of Western tuning systems and compositions can be rationalized by a theoretical model that describes a trade-off between minimizing sensory dissonance and maximizing compositional variety. From the Renaissance period onwards there was a trend towards more dissonance, and more compositional variety in both tuning systems and compositions. While this historical progression has perhaps been described qualitatively elsewhere, this model provides a more precise quantitative description of the phenomenon. The validity and scope of this model ought to be tested further by comparing its predictions with empirical measurements of tuning systems in both Western and non-Western cultures, alongside predictions of other theories of scale evolution. In the hope of encouraging and facilitating more of these interdisciplinary endeavors, I discuss some of my anecdotal experiences as a physical scientist embedded in the music science community, and offer advice on how to achieve better understanding and communication across disciplines.

INTRODUCTION

A few years ago, Jesse Berezovsky (2019) produced a highly original study that used the tools of statistical physics to study the evolution of tuning systems and scales. In this issue of *Empirical Musicology Review*, he and his collaborators Ryan Buechele and Alex Cooke have produced a study where the focus has shifted from statistical physics to music theory and the historical progression of tuning systems in Western Europe (Buechele et al., 2024).

      I am one of a rather limited set of people who have researched both statistical physics and evolution of scales (McBride & Avendaño, 2017, 2018; McBride & Tlusty, 2020; McBride et al., 2023; Brown et al., 2024). As such, my goal with this commentary is to first summarize the work of Berezovsky in a way that may be easier to digest for non-physicists. That's not to say that the original work is opaque, but given the diversity of the prospective audience, we need to approach communication from many angles. Through discussions with music scientists (I use 'music scientist' to denote anyone who studies music using the scientific method), it is clear to me that many people struggled to grasp the details of the original paper (Berezovsky, 2019). Thus, the new paper, almost entirely stripped of equations, is a welcome addition that I believe is approachable to a wider audience (Buechele et al., 2024). My summary aims for the middle ground: the music scientists who want mathematical detail. I plan to explain the models that Berezovsky uses, and to highlight how they complement the historical evolution of Western tonality.

      I have also spent the past five or so years conducting a comparative study on the evolution of scales, which has given me a good vantage point to put the work in a broader context. The evolution of tonality is quite a complex problem [2]. Scales can take on many forms, ranging from the ephemeral (a person singing to themselves; difficult to measure [3]) to the exceedingly robust (mathematical descriptions of scales, which have lasted millennia). As such, a theory of the evolution of vocal scales sung in prehistoric societies may not bear much relevance for the more mathematical scales of Western Europe in the common-practice period. It is with this contrast in mind that I will try to highlight where the work of Berezovsky fits.

      In this article, I summarize and critique the paper, and follow this up with a commentary on the art of interdisciplinary science and communication. The inscrutability of the original article to those in the field of music

science leads to the question, who was the intended audience? Physicists, or music scientists? Is it possible to write for both at the same time? I have struggled with this in my work, as I find that people regularly misunderstand my research in fundamental ways. In essence, I (we) have a communication problem. After discussing with several other scientists who have crossed the bridge from the physical sciences to study music, I have found that the experiences I have encountered are quite common. I will put forth some salient differences that I've noticed between the physical science and music science communities, in the hope that each side can better understand the other. Finally, I will offer some advice to both sides, so that common pitfalls may be avoided. Primarily, I think the route to better communication is *more* communication. The inclusion of physical scientists in music science can lead to serendipitous synergy, and hopefully this commentary can pave a way forward.

To conclude this introduction I look to Hermann von Helmholtz, a physicist and psychologist, and a pioneer of music science. From the introduction (translated by Alexander Ellis) of his groundbreaking work, *Sensations of Tone* (1885), I provide a quote that is still as relevant today as it was then:

> "In the present work an attempt will be made to connect the boundaries of two sciences, which, although drawn towards each other by many natural affinities, have hitherto remained practically distinct – I mean the boundaries of *physical and physiological acoustics* on one side, and of *musical science and esthetics* on the other. The class of readers addressed will, consequently, have had very different cultivation, and will be affected by very different interests. … I hope to have consulted the interests of both classes of readers." (p. 1)

SUMMARY OF BUECHELE ET AL. (2024)

The Question

The model of Berezovsky (2019) answers the question: "What is the optimal distribution of pitches within an octave that minimizes dissonance and maximizes compositional variety?" The model makes a key assumption that all pitches are equally likely to interact (form harmonies) with all other pitches. Dissonance, *D*, is quantified using the sensory dissonance model of Sethares (1998), while compositional variety is represented as the entropy, *S*, of the pitch distribution (lower entropy means fewer options for a composer). These two factors act against each other, as excluding dissonant pitch combinations leads to fewer options for the composer. The balance between them is controlled by a parameter, referred to as 'temperature', *T*, by analogy with statistical physics. The optimum pitch distribution is the one that minimizes *F* (analogous to the 'free energy' in statistical physics),

$$F = D_{tot} - TS \,, \tag{1}$$

where D_{tot} is the total dissonance of the tuning system (see Eq. (5) for the mathematical definition).

Quantifying Dissonance

As mentioned, the model chosen for quantifying dissonance is the sensory dissonance model of Sethares (1998). This model builds on the finding that people can reliably discern 'beats' (McDermott et al., 2016; Harrison & Pearce, 2020; Milne et al., 2023). 'Beats' occur when the interference pattern produced between tones with similar frequencies – e.g., tones that are one semitone apart and occur simultaneously – is perceptible (Sethares, 1998). This phenomenon is also observed for tones differing by a major seventh, which leads to the conclusion that the overall sensory dissonance between two complex tones – tones that consist of multiple, overlapping pure tones – can be computed by calculating interference between all pairs of overtones.

Dissonance, according to the model, depends on not only the frequency ratio of two complex tones, but also on the absolute frequency (via a term called the critical bandwidth, w_c) and the timbre (the relative weight of different overtones). Berezovsky (2019) uses a critical bandwidth of $w_c = 0.03$ and a sawtooth timbre. The dissonance, *d*, between two pure tones f_1 and f_2 is

$$d(f_1, f_2) = e^{-[ln(|\Delta x|/w_c)]^2}/w_c, \tag{2}$$

where $\Delta x = log_2(f_1/f_2)$ is the log-frequency difference, measured in units of octaves (this can be multiplied by 1200 to get units of cents). The dissonance D between two complex tones is obtained by summing d between pairs of overtones over an infinite series of overtones,

$$D(f_1, f_2) = \textstyle\sum_{n,m} l_{nm} d(nf_1, mf_2),\tag{3}$$

where n and m are terms in an integer series; $l_{nm} = min(n, m)^{0.606}$ approximates the perceived 'loudness' of the pair of nth and m^{th} overtones, as defined by the sawtooth timbre. In practice, one does not need an infinite series as higher overtones have diminishing contributions; summing over the first 10 overtones of each complex tone is likely to be sufficient.

To test the generality of these results, one could try to re-do this study with a different set of model parameters (w_c, or timbre), or indeed a different model (Harrison & Pearce, 2020). One could even replace sensory dissonance with a model based on harmonicity, or a composite model modeling both sensory dissonance or harmonicity (Marjieh et al., 2024). However, there is a good chance that the results will be similar, since the outputs of different models for sensory dissonance are strongly correlated for harmonic complex tones (McBride & Tlusty, 2020). The authors are aware of these possibilities and note that the reason for their choice was *simplicity*, following "the tradition of the physics community" (Buechele et al., 2024, p. 122).

Compositional Variety

In Buechele et al. (2024) the authors use the term compositional variety to describe the extent of possible choices for pitches that can be used in composition. Pitch is, of course, only one way among many of achieving variety in composition, but it is the relevant factor when considering scales. The compositional variety is modeled as the entropy of the pitch class distribution $P(x)$,

$$S = -\int_0^1 P(x) \, log(P(x)) \, dx,\tag{4}$$

where the integral is taken over $0 \leq x \leq 1$ octaves. This is a fairly simple concept: entropy is maximum when all pitches are equally likely; entropy is zero when there is only one choice for pitch. At very high temperatures ($T > 21$); when high compositional variety is favored) we find a uniform distribution over all pitches, which does not describe many (if any) types of music. As the temperature is lowered ($16.6 < T < 21$), discrete pitch classes emerge, starting with 12-tone equal temperament (12-TET). At this stage, there is no clear tonal hierarchy, so the distribution describes music that might be created using twelve-tone serialism. At still lower temperatures ($T < 16.6$) we see pitch distributions with tonal hierarchies that are more characteristic of Western classical music (Harasim et al., 2021). Eventually at $T = 5$, the lack of dissonance associated with unison or octave harmony leads to a single pitch class.

The Solution

The model describes pitch using a probability distribution $P(x)$ where the pitch class x is collapsed onto a single octave range, $0 < x < 1$ octave. The overall dissonance D_{tot} is the probability of two pitch classes, x and y, sounding at the same time $P(x, y)$ times the dissonance caused by the interval $x - y$. We can convert x and y, measured in octaves, into frequency using $f_1 = f_0 2^x$ and $f_2 = f_0 2^y$, by taking any real positive value for f_0; the choice of f_0 does not matter since absolute pitch is taken into account separately through the choice of the critical bandwidth, w_c. This allows us to calculate the overall dissonance as a function of x and y,

$$D_{tot} = \int_0^1 \int_0^1 P(x, y) D(x, y) \, dxdy.\tag{5}$$

The problem is simplified by assuming that two simultaneous pitch classes are independent of each other (in other words, all harmonies are equally likely), so that $P(x, y) = P(x)P(y)$, and

$$D_{tot} = \int_0^1 \int_0^1 P(x)P(y)D(x, y) \, dxdy.\tag{6}$$

The resulting optimal pitch class distribution $P(x)$ is the one that minimizes

$$F = D_{tot} - TS, \tag{7}$$

and after substituting Eq. 4 and Eq. 6 into Eq. 7,

$$F = \int_0^1 \int_0^1 P(x)P(y)D(x,y) \ dxdy + T\int_0^1 P(x) \ log(P(x)) \ dx. \tag{8}$$

Solving Eq. 8 to get the optimal $P(x)$ is not trivial, hence, the important innovation of Berezovsky is to notice that this is equivalent to a mathematical equation that was already solved in the field of statistical physics (Wagner, 1951).

Relaxing assumptions on how pitches interact

The assumption that simultaneous pitches are independent (i.e., all harmonies are equally likely) is of course incorrect, and it was made to simplify the problem. In reality, certain harmonies are more likely than others within a musical tradition (Jordania, 2011; Huang et al., 2017), and this is acknowledged by the authors. For example, if one studies the statistical distribution of pitch in Bach's Well-Tempered Clavier, one can find close to an even distribution of pitches overall, but minor second harmonies will occur less frequently than perfect fifths. Although not directly addressed in Buechele et al. (2024), Berezovsky offers the tone lattice model as an alternative formulation, where pitches only interact with a subset of neighbors (only few harmonies are likely) in an abstract 2-dimensional (or in later work, 3-dimensional) space (Berezovsky, 2019; Din & Berezovsky, 2023). The 2-dimensional tone lattice models lead, interestingly, to equidistant 5- and 7-note scales, which are surprisingly common across cultures (McBride et al., 2023); these scales are also the only equidistant scales with fewer than 12 notes which can accommodate a maximum number of 'imperfect fifths' (fifths, but allowing for some deviation from a frequency ratio of 3:2) (McBride & Tlusty, 2020). The 3-dimensional tone lattice models lead to tonal hierarchies that approximate the major/minor tonal hierarchies of Western classical music. While the original assumption of independent pitches was an oversimplification, these additional models show how one can gradually add complexity to a model to arrive at more detailed predictions that accord better with empirical data.

A new view on an old question

It is well known that one can arrive at a version (Pythagorean) of the Western tuning system via the circle of fifths. Starting from any note, progressively increment up and/or down in intervals of a fifth (Barbour, 1951) and you end up with something resembling 12-TET (or just intonation, mean-tone temperament, etc.), such that differences between scale degrees are typically smaller than a Pythagorean comma (20 cents). The approach of Berezovsky shows that you can achieve the same result, but from a different starting point: as a trade-off between minimizing sensory dissonance and maximizing compositional variety.

Comparison with historical evolution of Western intonation

The new component of Buechele et al. (2024) compared to Berezovsky (2019) is the comparison of the results of the model with the evolution of tonality in Western classical music. The authors first compare various tuning systems (equal temperament, mean-tone temperament, just intonation) with optimal tuning systems obtained at different values of T. This suggests that the historical progression of Western tuning systems can be explained as increasing compositional variety while simultaneously increasing sensory dissonance (increasing T). They next describe the evolution of Western tonality by depicting pitch distributions onto a radial plot based on the circle of fifths, showing results consistent with other investigations (Huang et al., 2017; Moss et al., 2023). This describes the gradual evolution of complexity of Western classical music, and they map this onto their model, showing that this is equivalent to music having a higher temperature (there was a shift towards greater compositional variety). They point out that in the 20th century, Western art music became increasingly dissonant, and some composers even eschewed scales entirely to compose using continuous pitch. This account offers a quantitative view of the evolution of Western classical music as a shift in the balance between minimizing sensory dissonance and maximizing compositional variety. One could say that this historical progression has certainly been stated before by musicologists, if you accept a comparison between compositional variety and the ability to easily play in many keys on one instrument (Barbour, 1951; Christensen, 2002). Nonetheless, it is useful to also have a more precise mathematical description of this progression, to supplement the qualitative accounts of musicologists and historians.

WHERE DOES THE WORK FIT INTO
WHAT WE KNOW ABOUT SCALE EVOLUTION?

Having first summarized and interpreted the work, I would like to examine where it fits in with the bigger picture of scale evolution. In the following I discuss: the factors that may have contributed to scale evolution; how compositional variety and dissonance fit into this picture; and the implications of pitch variability on the predictions of differences between theoretical Western tuning systems.

How did scales evolve?

The evolution of scales over time is a cultural-evolutionary process (Youngblood et al., 2023), and probably the most salient property of scales is that they are relatively fixed within a society, across performers and across performances (Surjodiningrat et al., 1972; Church, 2015; D'Amario et al., 2020; McBride et al., 2023) – although sometimes more than others (Arom et al., 2007; Weisser & Demolin, 2013; McBride et al., 2023). The question of evolution then becomes a question of how the scales change over time, and how some scales end up becoming used more frequently than others. Instead of modeling the cultural-evolutionary dynamics (Aucouturier, 2008), a more common approach is to study the properties of historical and extant scales (Huron, 1994; Gill & Purves, 2009; McBride & Tlusty, 2020; Phillips & Brown, 2022a, 2022b).

By studying common scales, if they have common properties (e.g., common harmonic intervals), then a plausible theory is one that links those properties to some evolutionary advantage (in the sense that the scales are more likely to persist or be selected). Other properties besides sensory dissonance and compositional variety have been investigated: Harmonicity has been considered instead of sensory dissonance, however the two properties are highly correlated so will not produce very different results (Gill & Purves, 2009). The interval spacing theory proposes that intervals need to be large enough that they are easily communicated without errors (Pfordresher & Brown, 2017; Phillips & Brown, 2022a, 2022b). Some have proposed that scales should be 'compressible' (few unique interval sizes), which leads to less complex melodies (McBride & Tlusty, 2020); others have proposed the opposite - that scales should have the properties 'uniqueness' and 'hierarchizability', where scales have more unique interval sizes (Balzano, 1980; Verosky, 2017). One important difficulty in studying the evolution of scales is that scales likely evolved differently in different cultures; e.g., societies with only solo monophonic music may not care about harmony (McDermott et al., 2016).To really understand how scales have evolved, one needs to study many properties, and many cultures.

I want to highlight how alternative hypotheses can lead to similar predictions as those found in the work of Berezovsky. The model of Berezovsky leads to separations between neighboring tones of at least about 1 semitone, since sensory dissonance is maximum when two tones that are very close in frequency are heard simultaneously. This is indeed found to be true in most societies, as intervals smaller than a semitone are exceedingly rare (Phillips & Brown, 2022b; McBride et al., 2023; Brown et al., 2024). The interval spacing hypothesis is a basis for a plausible alternative explanation of this phenomenon.

How much compositional variety is required? How can it be achieved?

The scope of compositional variety in Buechele et al. (2024) is limited to tonality. It is true that increasing the range of possible tones increases the number of possible compositions. It is worth pointing out, however, that one does not need a lot of tones in order to create varied music. Many societies use tonal systems with as few as two or three notes (Brandel, 1961; Kunst, 1967). Due to combinatorial explosion, even two note scales are sufficient to generate huge variety: the number of unique melodies can be calculated as N^L, where N is the number of possible tones, and L is the length of a melody. For $N = 2$ and length $L = 40$, there are roughly a trillion ($2^{40} \sim 10^{12}$) possible melodies. Moreover, pitch is only one component of music. Compositional variety can be widened through increasing complexity in rhythm, harmony, timbre, tempo and volume. Taking these into account, it is clear that there is no dearth of possible combinations to compose music with.

Nonetheless, there is indeed evidence in modern art music of composers becoming unsatisfied with the sets of tonalities available to them; the authors note examples from Debussy and Schoenberg, to Pousseur and Xenakis (Buechele et al., 2024). In summary, there is evidence of composers breaking the bounds of conventional Western tonality to increase compositional variety, but it is not clear how general this is for theories of the evolution of tonality, as it is possible to achieve compositional variety without resorting to increasingly dissonant tones. Furthermore, it

remains to be established whether a drive towards increased compositional variety is found outside of Western (or other) 'classical' musics, compared to stabilizing forces in music (Youngblood, 2019).

Dissonance as a determinant of tuning systems

The model in Berezovsky (2019) and Buechele et al. (2024) assumes octave equivalence and that scales span an octave. There is limited evidence to show that octave equivalence may not be universal (Jacoby et al., 2019), but there is much evidence of melodies that do not always span an octave (Brandel, 1961; Kunst, 1967; McBride et al., 2023; Brown et al., 2024). The model also assumes that dissonance is described by the dissonance curves of Sethares (1998), which compute the overall dissonance of complex tones as the sum of dissonance of the partial harmonics. Sensory dissonance of two pure tones is often found to lead to aversion (McDermott et al., 2016; Harrison & Pearce, 2020). However, the universality of dislike of sensory dissonance arising from interactions between overtones is hotly debated [4]. Third, the model assumes that dissonance is something to be minimized (as far as allowed by limits on compositional variety). There are differences between societies in terms of how dissonance is used in composition, which is independent of how much compositional variety there is. For example, both Western religious choruses and Lithuanian *sutartinės* (Ambrazevičius, 2017) tend to use scales with seven notes, yet sensory dissonance is avoided in the former and an integral part of the latter. Also, Javanese and Balinese gamelan music are highly similar, as expected given their shared history, yet only in Bali do they tune pairs of instruments with slight differences to create beats (Spiller et al., 2022). In summary, while there is evidence of preferences for harmony that avoids sensory dissonance in many cultures, it is not clear that dissonance is always something to be minimized.

Differences between temperaments… to whom does it matter?

Considerable attention is paid by the authors to differences between tuning systems (meantone temperament, just intonation, 12-TET) that were employed in the history of Western art music. Here I ask whether these differences are reliably produced and heard, and by whom?

Human pitch perception has clear limitations: Just-noticeable differences (JND), the smallest difference in pitch that can be reliably perceived, can be as low as 2-5 cents in laboratory conditions (Moore, 1973), which would enable distinctions to be made between tuning systems. However, JNDs are much higher for amateur musicians and non-musicians, and are much higher for interval differences compared to pitch differences (McDermott et al., 2010; Zarate et al., 2012). Interval discrimination, which is more relevant for tuning perception, is also likely to be diminished in realistic settings in comparison to controlled experiments in laboratories. Pitch production is also limited in precision, especially when it comes to singing (Ternström & Sundberg, 1988; Devaney et al., 2011; Pfordresher, 2022). Juxtapose the limits of pitch perception with the limitations of human vocal precision and it seems that the differences in Western tuning systems (typically about 10 cents) are unlikely to be reliably differentiated in practice, as has been found empirically (Barbour, 1951; Hagerman & Sundberg, 1980; Kopiez, 2003; D'Amario et al., 2020).

Despite these arguments against the ability to produce and perceive differences between similar tuning systems, we must take into account the historical progression of Western tuning systems, as noted in Buechele et al. (2024). The scholarly record dictates that tuning systems evolved from Pythagorean tuning, to just-intonation, mean-tone temperament, and settled at equal temperament. This indicates that someone cares. However, one could ask, who cares? I speculate that it is primarily the people who professionally create and tune fixed-pitch instruments. These are the people who create devices such as monochords and tuning forks, which improves the precision and stability of the intonation of fixed-pitch instruments over time. What we know of the historical development of Western tuning is obtained through the treatises of theoreticians, but due to the evanescence of musical sound we know little of how well theory corresponded to practice (Christensen, 2002). In modern times there is certainly evidence of a disconnect between theorists, who prefer simplicity and mathematics, and practitioners of Thai classical music (Garzoli, 2020). A better understanding of this is required to truly know how music evolution is differentially driven by performers, tuners, and theorists.

Summary

The work of Buechele et al. (2024) shows how the optimal (according to their theory) tuning systems change along a continuum from the least dissonant, to tuning systems that enable maximal compositional variety. They show that the evolution of tuning systems in Western classical music is commensurate with movement along this optimum, in the direction of increasing compositional variety. Furthermore, this is mirrored in the evolution of tonality in the work of

Western classical composers. A conservative estimate of where this model is applicable is in describing the evolution of Western classical music, at least in the precise tuning of instruments and the dissonance of compositions.

There are three areas that the work does not touch on, which I think would be extremely valuable next steps: (i) Are the small differences between theoretical tuning systems (e.g., just intonation vs. 12-TET) evidenced in empirical measurements of tunings or performances? One can measure intonation of vocal performances or how instruments are tuned; the differences between tuning systems may be negligible compared to the empirical variance in intonation. Existing studies so far lean towards this result, but there are only a few such measurements (Hagerman & Sundberg, 1980; Kopiez, 2003; D'Amario et al., 2020). (ii) How does the proposed theory compare with other theories of the evolution of music? (iii) How well do the results describe tuning systems from other cultures? This last question can be evaluated with recently-published datasets of scales (McBride et al., 2023; Brown et al., 2024). I anticipate future work that leads to clarification of the relevance and power of this theory in explaining the evolution of scales more generally.

Although this Commentary was invited in response to Buechele et al. (2024), it largely deals with the theoretical model of Berezovsky (2019). This particular mode of scientific communication – whereby much of the same information is presented in two separate, stylistically-distinct papers to two very different audiences – is uncommon (though not unheard of). What would Hermann von Helmholz have chosen to do in this case, had he lived today? It is with this in mind that I write on the more general issue of interdisciplinary communication.

THOUGHTS ON INTERDISCIPLINARY COMMUNICATION

I would like to start the following discussion with a disclaimer. I will here be indulging myself in the writing of opinion, based on anecdotal evidence and conjecture. Some of my opinions are likely wrong, probably due to under-sampling of experiences, and possibly due to faulty inference. Hence, I feel the need to begin by introducing myself. I received education in chemical engineering and statistical physics, and after my PhD I shifted to biophysics. As luck would have it, I ended up with sufficient academic freedom to also study music. And so, I went down the path of studying various research questions, attending conferences, and I developed a network of music scientists to discuss and collaborate with.

When I was first invited to write a commentary on Buechele et al. (2024), I knew I wanted to speak about the general problem of interdisciplinary communication. To supplement my own experience, I sought out stories from others who have also straddled the boundary between physical and music sciences. I was surprised by the extent of overlap between different researchers' experiences, which gives me hope that this discussion is a worthy endeavor after all [4]. In the following, I discuss the importance of interdisciplinary work in music research, epistemological differences between disciplines, and advice for both physical scientists and those more fully embedded in the music science community.

Music science is about as interdisciplinary as it gets. Beyond the core disciplines, such as music theory, ethnomusicology, historical musicology and comparative (systematic) musicology, there are contributions from psychology, neuroscience, computer science, mathematics, anthropology, archaeology, and cultural evolution. Given this diversity, it is not surprising that there are physical scientists who also want to join in the fun. What do they bring to the party? Buechele, Cooke, and Berezovsky (2024) show how one can repurpose existing solutions from physics if two models share the same mathematical forms. Others have used theory and tools from biology to study the evolution of music: evolutionary theory (Nakamura & Kaneko, 2019; Lambert et al., 2020), sequence alignments (Savage et al., 2022), and morphometrics (Chitwood, 2014; Aguirre-Fernández et al., 2020, 2021). I should also point out that such transfer of ideas is not unidirectional, as tools developed for studying music have been used in the physical sciences (Leroi et al., 2020; Zali et al., 2023).

To reap the benefits of interdisciplinary research, we need effective communication and collaboration, followed by constructive commentary and criticism. Given the heavy burden on academics (research, teaching and administrative jobs; the constant drive to write papers and grants), stepping outside of one's main field is not necessarily supported at the faculty level. This means that people who cross the boundary from physical sciences to music, do so sporadically, and do not have so much time to commit to research and networking outside their principal domain. For these reasons, it is especially important that we make the most of what few opportunities arise. With this in mind, I would like to offer some advice to both sides on how to better communicate and understand each other.

Observations on differences between physical sciences and music science

I tentatively start this section with a conjecture, that differences in philosophy between disciplines can arise due to differences in the degree of certainty one can have in scientific results of that discipline. The lack of certainty in results

is of course related to the degree of predictability and manipulability (Weaver, 1948; Sanbonmatsu et al., 2021) in the subject (e.g., psychology and culture are typically less predictable than inert matter), so this is in no way meant to be a criticism of the representatives of a particular discipline. This conjecture has been stated elsewhere more eloquently, such as in discussions on the hypothesized "hierarchy of the sciences" (Comte, 1855; Fanelli & Glänzel, 2013; Simonton, 2015). On one end of the extreme is mathematics, where once certain axioms have been assumed, there is a correct answer that can be unambiguously arrived at. On the other end of the extreme is a discipline like economics, where even after assumptions (which are not as clear as mathematical axioms) have been laid down, experimental results can be questioned based on sampling biases, construct validity, insufficient statistics, and the fidelity or utility of models used (Eronen & Bringmann, 2021). Disagreements are rare in the former discipline (Lamers et al., 2021), while in the latter disagreements have historically been embedded in the system, in the form of different schools of thought. I have noticed in my experience (and the experience of others) a series of distinctions (distinctions of degree, not categorical distinctions) that I would like to summarize: (i) the weight that is given to argument and logic over empirical results; (ii) tendency for debates to become polarized; and (iii) how much evidence is needed to claim 'proof'.

LOGIC OVER EMPIRICISM

Logic is an essential tool in all sciences. It seems to me, however, that logic and argument are given greater weight in music science than in the physical sciences. I think that this arises naturally when there is greater uncertainty in observations - for example, the only significant scientific arguments I have observed in the physical sciences is when there is a challenging technique which has not undergone sufficient testing. The fact is that many observations in music science are statistical, such that there are few infallible 'laws'. There are few such irrefutable starting points (e.g., the physics of sound, Weber-Fechner laws) from which to build upon. With this amount of uncertainty, observations, inferences, and deductions can all be questioned at any level of a research project. I think it is this uncertainty that leads to increased reliance on argument and logic (Deroover et al., 2023).

As an illustration, I ask the reader to consider a recent pair of target papers in the journal *Behavioral and Brain Sciences* which each proposed a theory for the origins of music (Mehr et al., 2021; Savage et al., 2021). These articles provide two examples that would have been unlikely to occur in the physical sciences: First, in one paper the authors (Mehr et al., 2021) make an *extensive seven point qualitative argument* against the theory of music as an exaptation. Qualitative arguments are not unheard of in physical sciences, but what stands out here is the use of numerosity of arguments in place of quality (Harrison & Seale, 2021). Second, the papers are accompanied by *60 different commentaries* that each discuss an aspect raised (or not raised) in the target papers. A potentially comparable case in the physical sciences would be the origin of life, but there scientists are focused almost exclusively on exploratory work, generating empirical results, rather than debates over theories; perhaps there is a general acceptance of the fact that debate cannot currently solve the question. In drawing attention to these examples, I am not making any judgments on what is 'correct'. I merely want to use these examples to illustrate differences between scientific communities. For me, awareness of the differences in approach helps to understand researchers from different disciplines.

POLARIZATION

It is my impression that there is a greater degree of polarization within music science, compared to physical sciences. In discussions at music conferences, I have been asked what side I take on some dichotomy [5], while I never experienced this at physical science conferences. An irregular experience of mine (yet which is new to me since getting involved in music science) is that music scientists erroneously assume or deduce my thoughts on topics based on what I have said previously. The only explanation I can think of for this is that my previous statements labeled me as part of a particular 'tribe' (Becher & Trowler, 2001). Such community-wide disagreements within a discipline appear to me to be uncommon in modern physics [6]. They are more common in biology [7], but these disagreements typically get resolved after some decades due to methodological and technological advances. In contrast, one can point to the centuries-old nature-nurture debate over why music sounds pleasant [8]. This debate is today extremely active, and shows no sign of imminent resolution (Collins, 1994). I do not mean to imply that polarization is unique or exceptional in music science communities, but it appears to me to be more prominent than in the physical sciences.

WHAT CONSTITUTES PROOF?

Different disciplines vary in standards about how much evidence constitutes proof (i.e., generally accepted that something is probably true). High-energy particle physics is an extreme example, where the gold standard is six sigma, which means that a result ought to occur by chance less than one out of half a billion times ($p = .00005$) to be accepted. In contrast, the typical probability needed to refute a null hypothesis in psychology is one in twenty ($p = .05$). In general, results in physical sciences are not accepted widely until there are multiple repeat measurements, using multiple experimental techniques. While similar approaches are found in music science, I can think of one example that illustrates a lower threshold for acceptance: In 1980 it was proposed that musical scales which have a property termed *uniqueness* - where each tone has unique relations with other tones, in terms of interval sizes - have certain benefits (Balzano, 1980). Scales with this property are *asymmetric*, and it has been speculated that this may facilitate key finding, increase the potential of melodic variation, and allow tones to assume different functions (Trehub, 1999). Despite going untested for almost twenty years, Balzano's (1980) paper racked up 110 citations. After the publication of a *single* experimental investigation (Trehub et al., 1999) (I have nothing bad to say about the paper itself), the 1980 paper picked up another 300 citations in the next 20 or so years. Eventually a more thorough investigation was undertaken (Pelofi & Farbood, 2021), but even before this I had spoken to many people who seemed to take this as 'established knowledge' in the field, yet it consisted of a hypothesis with a single test (participant sample size < 40). It is unlikely that such a paper would have been as influential in physics, where hypotheses only become important when they explain or predict robust observations [9].

Once again, I need to clarify that my intention is not to criticize, but to highlight differences between disciplines. Furthermore, one should be aware that these differences may exist, in part, due to the relative "cost" of psychological vs physical experiments. Physics projects can burn through a great deal of money, but they are less constrained by legal and ethical concerns (Sanbonmatsu et al., 2021), which severely limit the types of psychological experiments that are feasible, and how many people can be recruited. Regardless, I do worry that the inferred lower burden of proof in music science may lead to overly stringent gatekeeping, as a single peer-reviewed publication appears to carry much more weight than it does in physical sciences. I would like to illustrate this with a personal story: In a discussion about peer-review with a music scientist, my colleague opined on the importance of rejecting papers lest they lead the community astray. I explained that the papers and analyses could speak for themselves, and that people would figure it out. I was laughed at (in good humor) for my 'innocence'. I don't know if it was due to my (rather youthful, still) age, personality, or the standards of proof that I learned through the culture of physical sciences. Nonetheless, I wonder how different standards of proof may lead to different behaviors in academia.

In this section I have endeavored to open a discussion on the differences between disciplines, based largely on my personal experience. While readers may agree or disagree with my various conjectures, inferences, and deductions, I hope that they realize that they are made with good intent. In the following I aim to elaborate on more practical advice for both sides of the continuum, with the hope that the advice facilitates productive interactions between scientists differentiated by background, but united in their curiosity for the wonders of music.

Advice on Communicating Across the Divide: For Physical Scientists

DISCUSS & COLLABORATE

It may seem obvious to the reader that more discussion and collaboration is beneficial. Here I would like to delve into the nature of the benefits, and how to practically achieve more discussion and collaboration. For me, the main contribution of a music scientist is the breadth and depth of knowledge in their subject. They know answers to many important questions that are necessary to guide reductive approaches, so that our models do not have glaring mistakes and rely on baffling assumptions. They know what has been tried before, and what are the common pitfalls within a topic. Expertise (especially the sort that is not directly evident in published papers) is invaluable in developing psychological experiments and performing appropriately-controlled big-data analyses. Some of the most enlightening discussions I have had have been with those whose expertise overlaps the least with my own. It is through these discussions that I have also learned something about the broad community of music science, which has led me to become a more effective communicator in papers and presentations. Finally, I speculate that papers may be more widely read if they are written by a diverse set of co-authors (Galam, 2022), given the social nature of knowledge diffusion (Wallace et al., 2012).

The more useful question to answer is: "How to find people to discuss and collaborate with?" I started out mainly with cold emails to strangers; this strategy has a variable hit-rate, but costs very little. Eventually I became aware of the many conferences and meetings (both online and offline), where I have found people to be

overwhelmingly welcoming and certainly interested in speaking to a physical scientist. In my experience, the physical sciences have not yet adapted to the new world of online meetings - platforms are basic, buggy, and interactions have been limited. In contrast, the online music conferences I have attended have made attempts to be inclusive of different time zones, make use of innovative platforms like Gathertown, asynchronous interactions through platforms like Slack, and overall people seem more engaged. To help make these conferences and meetings more visible, I have compiled a list hosted on Github (https://github.com/jomimc/MusicScienceMeetings), which I hope can be maintained over time.

The last issue I would like to address is the question of how to initiate collaborations. This is difficult by default, since for a physical scientist with an established network in their main field, it requires either going outside one's network. This is particularly daunting since the usual development of an academic network starts with academic colleagues and advisors, and then organically grows throughout postgraduate study. Starting a new network from scratch requires a different approach, and some courage. It can also be difficult to start a collaboration when you have no track record of publishing in a different field. First, I advise starting with attending meetings and emailing people for discussions – in my experience people are more than happy to talk about their work if they have time. Second, I advise producing some preliminary results, and then to ask a music scientist for help in guiding the direction of the research at this early stage. One should be aware that such a preliminary study may suffer from serious design flaws and need to be entirely revised, but this shows a level of intent and commitment that I believe will help people take you seriously.

EFFECTIVE COMMUNICATION

A succinct summary for this section is: "Write longer papers!" I was taught that good writing is about concision and clarity, but the level of detail required to achieve clarity depends on your audience.

1. **State your assumptions clearly.** And also state what you do not assume. Too often I have found myself perplexed at being criticized for something that I did not state, nor think. I speculated earlier that this may be due to polarization / tribalism within sub-fields, but who really knows. In any case, it pays to play safe.
2. **Explain methods in greater detail.** The interdisciplinary audience of music science has a great diversity of expertise in quantitative and qualitative methods. As a result, it is prudent to carefully explain concepts at multiple levels of abstraction, and not to expect that people can easily follow just because everything is defined mathematically. One approach is to write two separate papers (Berezovsky, 2019; Buechele et al., 2024), while another approach is to write qualitatively in the main text with the mathematics in the appendix (Helmholtz, 1885).
3. **Be careful about the meanings of words**. Like in physics, in music science there are many commonplace terms that are used imprecisely by the layperson, which may have clear definitions within an academic field. Newcomers to the field ought to be aware of these definitions (this is where collaboration across disciplines can be of immense benefit). Moreover, some definitions are not even fixed within music theory since they differ across cultures. For example, I struggled in my own work on scales with how to use the term "mode": In the history of Western music theory it can refer simply to the set of pitches in a scale (Powers, 2001), but can also include the statistical distribution of pitches (Harasim et al., 2021) or some other form of hierarchy (Fernando, 2007). Analogous terms in non-Western traditions can be even more broadly defined, such that they also include information about transitions between pitches, the time-duration of specific pitches, and the use of specific ornaments (Akkoç, 2002; Church, 2015; Pearson, 2016; Ganguli & Rao, 2019).
4. **Spend more time explaining limitations**. As with assumptions, omitting discussions of caveats and limitations may be misconstrued as a lack of awareness of such issues.
5. **Draw analogies between disciplines.** Buechele et al. (2024) use the analogy of crystallization to explain the emergence of well-defined ordering between pitches in tonality. I believe it is pertinent to ask how this analogy aids understanding. This analogy certainly makes it easier for physicists to understand the work. Beyond that, the analogy seems poetic; inspiring, but perhaps not particularly informative. It requires a music scientist to learn something new (physics), and it does not give any new understanding of music other than that there is a mathematical correspondence with another phenomenon.

Advice on Communicating Across the Divide: For Music Scientists

The following advice is informed by discussions with several other physical scientists who have dabbled in, or migrated to music science. From these conversations, I have attempted to distill perceived shortcomings of critiques that we have received, to arrive at a concise set of concrete suggestions.

1. **Be kind, assume the best.** This may sound like a mere request for civility, but I also sincerely believe that it is useful advice in the pursuit of understanding. I think I can best illustrate this with my own personal story of writing the first part of this commentary. My initial response to Berezovsky (2019) and Buechele et al. (2024) was a laundry list of criticisms and caveats: sensory dissonance is not necessarily something to be avoided, compositional variety is not limited to intonation, etc. I decided to first have a chat with Jesse Berezovsky (in the spirit of good communication) to see what he actually thought. I was pleasantly surprised to find that he was aware of many of the things I had to say, and that we found little to disagree on. After some self-reflection, I realized that this is exactly the same process that I have often found when talking with music scientists. I have encountered many examples of a music scientist who aims to correct and inform me, and they also become pleasantly surprised after realizing that we mainly agree when the details are laid out. Thus, I advise against filling in the blanks (in a paper, or talk) with assumptions about the author's or presenter's beliefs.

2. **If you do not understand, ask.** Most disagreements I've encountered in disseminating my work were due to misunderstandings, rather than actual points of difference (Weinberger & Bradley, 2020). It costs little to send an email for clarification, and in my experience, authors are quick and willing to respond to direct questions about their work.

3. **Do not make judgments based on the presence (or absence) of domain experts.** It is unfortunate that a recurrent theme in my discussions with other physical scientists is that we have received criticism for not collaborating with 'appropriate' experts. As I made clear earlier, I certainly do value the expertise of scientists from different domains. However, in the contexts in which such objections have often been made, the criticism of lacking a domain expert has seemed to stand on its own legs. By itself, this is an *ad hominem* criticism, since it pertains to the assumed expertise of the author, rather than to the quality of their work. Suggestions to include domain experts are entirely appropriate in general, but work should not be judged simply because of the presence or absence of domain experts.

4. **Pay attention to results, not just to semantic disagreements or speculation.** Physical scientists are typically focused on the methods and results presented in a paper, and the discussion (where conclusions are often overstated, and speculation is intended to guide future work) is taken with a pinch of salt. I'm aware of cases where research performed by physical scientists was critiqued based on its speculation aspects, instead of its results. [10] To me, these are unfortunate examples of missed opportunities to appreciate the value of interdisciplinary research. It is this kind of missed opportunity that makes me urge others to reconsider their initial impulses, as I did after discussing with Berezovsky.

SUMMARY

In this commentary I have attempted to bridge the communities of physical and music sciences. In the first section I summarized two papers (Berezovsky, 2019; Buechele et al., 2024) which are like two sides of the same coin. One was written for physicists, and the other was written for musicologists. My summary is directed towards the middle ground. In the second section I tried to distill some advice and understanding from experience, with the aim of improving dialogue and comprehension between the groups. I spent much time highlighting differences, while emphasizing that these are differences of degree. On the other hand, I neglected to mention the substantial similarities, chief of which is categorical: music scientists and physical scientists are scientists foremost, who share an intense curiosity about the world and a drive to understand it using rigorous methods and reductive theories. As music is a joy for all mankind, I am sure that physical scientists will continue to be drawn to it. With humility and understanding, physical scientists and music scientists surely have much to share.

ACKNOWLEDGEMENTS

Chitwood, and Patrick Savage for comments on the manuscript. This article was copyedited by Gabriele Cecchetti and layout edited by Jonathan Tang.

NOTES

[1] Correspondence can be addressed to: Dr John M McBride, Center for Soft and Living Matter, Institute for Basic Science, South Korea, jmmcbride@protonmail.com

[2] The evolution of tonality is the result of complex cultural-evolutionary dynamics (Youngblood et al., 2023), constrained by biology (McBride et al., 2023). For example, the use of tonality in a society can change in response to the development of technology, influence from other cultures, and conquest (Church, 2015). To understand the evolution of tonality, not only must we understand human biology, but also historical context and contingency (Glenn, 1991).

[3] Coming from my own experience, musical scales are often extremely difficult to measure reliably. Instrumental music is typically easier (especially in the case of fixed-pitch instruments), but vocal music is a real bother (Brown et al., 2024). For a long time, the only rigorous measurements (this description excludes measurements by ear, notated on staff notation) of scales were obtained by measuring the tuning of instruments using tuning forks or a Stroboconn (McBride et al., 2023). Only more recently have there been quantitative measurements of scales from audio recordings of singing, and even today this is quite a time-intensive process (Rosenzweig et al., 2020).

[4] See (Chmutina et al., 2022) for a similar discussion of interdisciplinary issues in the field of disaster scholarship.

[5] I shall refrain from giving specific examples, since I don't want anyone to identify themselves in them. However, I can say that the areas where I have encountered signs of 'tribalism' are: the origins of consonance of intervals, the origins of musical scales, and the origins of music. There is a common theme here – these are all very deep questions that are difficult to answer.

[6] A famous example of a community-wide disagreement in physics is the one at the beginning of the 20th century over quantum theory (Marage & Wallenborn, 1999). More recently there has been a long-standing debate about how high-temperature superconductivity is possible, which has been especially acrimonious (Rodgers, 1998; Mann, 2011). In my academic career I've only witnessed one such disagreement, on the origins of some anomalous behavior of water, but what I saw was quite civil (Shi & Tanaka, 2020). Perhaps the difference that I've noticed is that in physics I've experienced more of an attitude of "let's wait and see how the results turn out", as opposed to battling over strong opinions.

[7] Tribalistic disputes appear to me to be more common in biology than in physics. In my own work on biophysics of proteins, I've encountered one example of tribalism in the literature over the hypothesized mechanisms of protein folding. Proteins are produced in an unfolded state, and spontaneously fold into a unique, intricate structure. The disagreement in this case was about whether they folded continuously, or in a two-state manner (a discontinuous transition between the unfolded and folded state), or via multi-state transitions. For an example, see this exchange of letters in the journal *Nature* (Ferguson et al., 2007; Sadqi et al., 2006, 2007; Zhou & Bai, 2007). The disagreement arose because of uncertainty in observations - there were only few proteins for which folding mechanisms had been reported, the resolution of experimental techniques was low, and experimental techniques were indirect, relying on multiple uncertain inferences.

[8] In music science, I've had the (guilty) pleasure of witnessing a recent dispute since I have started working in the field: The ongoing debate about whether perception of consonance owes more to learning or innate interactions between sound and the brain. In the last seven years, there have been four publications that have been rapidly followed by critical responses: one paper suggested that consonance is learned rather than innate (McDermott et al., 2016; Bowling et al., 2017; Savage & Currie, 2017); one paper suggested that consonance is due to vocal similarity (Bowling et al., 2018a; Goffinet, 2018; Bowling et al., 2018b); another paper supported an innate account of consonance, but not related to vocal similarity (Friedman et al., 2021a; Harrison, 2021; Smit & Milne, 2021; Friedman et al., 2021b); the most recent paper is a review that proposes auditory roughness as the most dominant determinant (Di Stefano et al., 2022; Lahdelma & Eerola, 2023; Bowling, 2023). It is worth noting that this question about the origins of consonance is *old* (Helmholtz, 1885). While I have had little trouble enumerating examples of tribalism in biology,

they all had limited life-spans, as sufficient evidence was accumulated to achieve some sort of resolution. I am skeptical about the possibility of imminent resolution on the question of whether consonance is more innate or learned. There is some hope, however, thanks to the increasing use of innovative approaches that increase the breadth and extent of both subject populations (Long et al., 2023), and collaborations (Ozaki et al., 2022).

[9] The hypothesis that *asymmetric* scales have musicological / cognitive functions is typically cited as an explanation for the observation that symmetric (equidistant) scales are rare. Somehow the lack of symmetry in scales has been something that people feel the need to explain, despite the fact that symmetry in general is rare - try counting all of the possible scales and see how many of them are symmetric. Add to this the fact that there are many examples of symmetric scales in the literature, yet rarely cited (Ross & Knight, 2019; McBride et al., 2023). Perhaps the reason people assumed that equidistant scales ought to be more common (by chance) is due to thinking about symmetry and asymmetry as two classes (and thus should be represented equally), without giving due regard to the fact that there are many more possible asymmetric scales than symmetric ones.

[10] There is an interesting paper that proposed a method that can quantify whether a time series (e.g., a melody) is reversible or not (González-Espinoza et al., 2020). In the case of music, if music was reversible, then it would have similar properties if heard in reverse (not in the sense of playing a Led Zeppelin LP backwards to hear satanic messages, but rather if the order of notes was reversed). It would be interesting to investigate the extent of time-reversibility in music cross-culturally, as this feature has been little explored to date. Unfortunately, the work was summarized in a commentary on Mehr et al. (2021) and Savage et al. (2021), not based on the main contribution of the work, but on the work's more exploratory discussion section (Margulis, 2021). Despite using phrases like "points to", and "possible relations", to indicate the authors' uncertainty, this speculative section was the focus of the commentary. While Margulis (2021) can be lauded for advocating interdisciplinary collaboration, I feel that such a commentary would be more effective at encouraging such collaborations if it included a more positive representation of González-Espinoza et al. (2020). For an example of a more positive response in a similar situation, see (Kulkarni et al., 2024; Cutts, 2024).

REFERENCES

Aguirre-Fernández, G., Barbieri, C., Graff, A., Pérez de Arce, J., Moreno, H., & Sánchez-Villagra, M. R. (2021). Cultural macroevolution of musical instruments in South America. *Humanities and Social Sciences Communications*, *8*(1), 1–12. https://doi.org/10.1057/s41599-021-00881-z

Aguirre-Fernández, G., Blasi, D. E., & Sánchez-Villagra, M. R. (2020). Panpipes as units of cultural analysis and dispersal. *Evolutionary Human Sciences*, *2*, e17. https://doi.org/10.1017/ehs.2020.15

Akkoç, C. (2002). Non-Deterministic Scales Used in Traditional Turkish Music. *Journal of New Music Research*, *31*(4), 285–293. https://doi.org/10.1076/jnmr.31.4.285.14169

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

Arom, S., Fernando, N., & Marandola, F. (2007, July 11). An Innovative Method for the Study of African Musical Scales: Cognitive and Technical Aspects. *Proceedings SMC'07*. 4th Sound and Music Computing Conference, Lefkada, Greece.

Aucouturier, J.-J. (2008). The Hypothesis of Self-Organization for Musical Tuning Systems. *Leonardo Music Journal*, *18*, 63–69. https://doi.org/10.1162/lmj.2008.18.63

Balzano, G. J. (1980). The Group-Theoretic Description of 12-Fold and Microtonal Pitch Systems. *Computer Music Journal*, *4*(4), 66–84. https://doi.org/10.2307/3679467

Barbour, J. M. (1951). *Tuning and temperament: A historical survey*. Michigan State College Press.

Becher, T., & Trowler, P. (2001). *Academic tribes and territories: Intellectual enquiry and the culture of disciplines* (2nd ed). Open University Press.

Berezovsky, J. (2019). The structure of musical harmony as an ordered phase of sound: A statistical mechanics approach to music theory. *Science Advances*, *5*(5), eaav8490. https://doi.org/10.1126/sciadv.aav8490

Bowling, D. L. (2023). Vocal similarity theory and the biology of musical tonality. *Physics of Life Reviews*, *46*, 46–51. https://doi.org/10.1016/j.plrev.2023.05.006

Bowling, D. L., Hoeschele, M., Gill, K. Z., & Fitch, W. T. (2017). The Nature and Nurture of Musical Consonance. *Music Perception*, *35*(1), 118–121. https://doi.org/10.1525/mp.2017.35.1.118

Bowling, D. L., Purves, D., & Gill, K. Z. (2018a). Reply to Goffinet: In consonance, old ideas die hard. *Proceedings of the National Academy of Sciences*, *115*(22). https://doi.org/10.1073/pnas.1805570115

Bowling, D. L., Purves, D., & Gill, K. Z. (2018b). Vocal similarity predicts the relative attraction of musical chords. *Proceedings of the National Academy of Sciences*, *115*(1), 216–221. https://doi.org/10.1073/pnas.1713206115

Brandel, R. (1961). *The Music of Central Africa.* Springer Netherlands. https://doi.org/10.1007/978-94-015-0997-8

Brown, S., Phillips, E., Husein, K., & McBride, J. (2024). *Musical scales optimize pitch spacing: A global analysis of traditional vocal music* [Preprint]. In Review. https://doi.org/10.21203/rs.3.rs-3928177/v1

Buechele, R., Cooke, A., & Berezovsky, J. (2024). Crystals of sound: Applying the physics of phase transitions to musical intonation. *Empirical Musicology Review*

Chitwood, D. H. (2014). Imitation, Genetic Lineages, and Time Influenced the Morphological Evolution of the Violin. *PLOS ONE*, *9*(10), e109229. https://doi.org/10.1371/journal.pone.0109229

Chmutina, K., Cheek, W., & von Meding, J. (2022). "Critique is not a verb": Is peer review stifling the dialogue in disaster scholarship? *Disaster Prevention and Management: An International Journal*, *31*(4), 387–397. https://doi.org/10.1108/DPM-09-2021-0266

Christensen, T. S. (2002). *The Cambridge history of Western music theory.* Cambridge University Press. https://doi.org/10.1017/CHOL9780521623711

Church, M. (Ed.). (2015). *The other classical musics: Fifteen great traditions.* The Boydell Press.

Collins, R. (1994). Why the social sciences won't become high-consensus, rapid-discovery science. *Sociological Forum*, *9*(2), 155–177. https://doi.org/10.1007/BF01476360

Comte, A. (1855). *The positive philosophy of Auguste Comte.* Blanchard.

Cutts, E. (2024, February 16). Secret Mathematical Patterns Revealed in Bach's Music. *Scientific American*. https://www.scientificamerican.com/article/secret-mathematical-patterns-revealed-in-bachs-music/

D'Amario, S., Howard, D. M., Daffern, H., & Pennill, N. (2020). A Longitudinal Study of Intonation in an a cappella Singing Quintet. *Journal of Voice*, *34*(1), 159.e13-159.e27. https://doi.org/10.1016/j.jvoice.2018.07.015

Deroover, K., Knight, S., Burke, P. F., & Bucher, T. (2023). Why do experts disagree? The development of a taxonomy. *Public Understanding of Science*, *32*(2), 224–246. https://doi.org/10.1177/09636625221110029

Devaney, J., Mandel, M. I., Ellis, D. P. W., & Fujinaga, I. (2011). Automatically extracting performance data from recordings of trained singers. *Psychomusicology: Music, Mind and Brain*, *21*(1–2), 108–136. https://doi.org/10.1037/h0094008

Di Stefano, N., Vuust, P., & Brattico, E. (2022). Consonance and dissonance perception. A critical review of the historical sources, multidisciplinary findings, and main hypotheses. *Physics of Life Reviews*, *43*, 273–304. https://doi.org/10.1016/j.plrev.2022.10.004

Din, H., & Berezovsky, J. (2023). Critical behavior and the Kibble-Zurek mechanism in a musical phase transition. *PLOS ONE*, *18*(1), e0280227. https://doi.org/10.1371/journal.pone.0280227

Eronen, M. I., & Bringmann, L. F. (2021). The Theory Crisis in Psychology: How to Move Forward. *Perspectives on Psychological Science*, *16*(4), 779–788. https://doi.org/10.1177/1745691620970586

Fanelli, D., & Glänzel, W. (2013). Bibliometric Evidence for a Hierarchy of the Sciences. *PLoS ONE*, *8*(6), e66938. https://doi.org/10.1371/journal.pone.0066938

Ferguson, N., Sharpe, T. D., Johnson, C. M., Schartau, P. J., & Fersht, A. R. (2007). Analysis of "downhill" protein folding. *Nature*, *445*(7129), E14–E15. https://doi.org/10.1038/nature05643

Fernando, N. (2007). Echelles et modes: Vers une typologie des systèmes scalaires. In J.-J. Nattiez (Ed.), *Musiques, une encyclopédie pour le XXIème siècle, vol. 5. L'unité de la musique* (pp. 945–979). Actes Sud.

Friedman, R. S., Kowalewski, D. A., Vuvan, D. T., & Neill, W. T. (2021a). Consonance Preferences Within an Unconventional Tuning System. *Music Perception*, *38*(3), 313–330. https://doi.org/10.1525/mp.2021.38.3.313

Friedman, R. S., Kowalewski, D. A., Vuvan, D. T., & Neill, W. T. (2021b). Response to Invited Commentaries on "Consonance Preferences Within an Unconventional Tuning System." *Music Perception*, *38*(3), 340–341. https://doi.org/10.1525/mp.2021.38.3.340

Galam, S. (2022). Physicists, non physical topics, and interdisciplinarity. *Frontiers in Physics*, *10*, 986782. https://doi.org/10.3389/fphy.2022.986782

Ganguli, K. K., & Rao, P. (2019). On the perception of raga motifs by trained musicians. *The Journal of the Acoustical Society of America*, *145*(4), 2418–2434. https://doi.org/10.1121/1.5097588

Garzoli, J. (2020). Competing Epistemologies of Tuning, Intonation and Melody in the Performance of Thai Classical Music on Non-Fixed-Pitch Instruments. *Sojourn: Journal of Social Issues in Southeast Asia*, *35*(3), 407–436. https://doi.org/10.1355/SJ35-3a

Gill, K. Z., & Purves, D. (2009). A Biological Rationale for Musical Scales. *PLoS ONE*, *4*(12), e8144. https://doi.org/10.1371/journal.pone.0008144

Glenn, S. S. (1991). Contingencies and metacontingencies: Relations among behavioral, cultural, and biological evolution. In P. A. Lamal (Ed.), *Behavioral analysis of societies and cultural practices*. Hemisphere Pub. Corp.

Goffinet, J. (2018). Little evidence for the vocal similarity hypothesis. *Proceedings of the National Academy of Sciences*, *115*(22). https://doi.org/10.1073/pnas.1804577115

González-Espinoza, A., Martínez-Mekler, G., & Lacasa, L. (2020). Arrow of time across five centuries of classical music. *Physical Review Research*, *2*(3), 033166. https://doi.org/10.1103/PhysRevResearch.2.033166

Hagerman, B., & Sundberg, J. (1980). Fundamental frequency adjustment in barbershop singing. *STL-QPSR*, *21*(1).

Harasim, D., Moss, F. C., Ramirez, M., & Rohrmeier, M. (2021). Exploring the foundations of tonality: Statistical cognitive modeling of modes in the history of Western classical music. *Humanities and Social Sciences Communications*, *8*(1), 5. https://doi.org/10.1057/s41599-020-00678-6

Harrison, P. M. C. (2021). Three Questions Concerning Consonance Perception. *Music Perception*, *38*(3), 337–339. https://doi.org/10.1525/mp.2021.38.3.337

Harrison, P. M. C., & Pearce, M. T. (2020). Simultaneous consonance in music perception and composition. *Psychological Review*, *127*(2), 216–244. https://doi.org/10.1037/rev0000169

Harrison, P. M. C., & Seale, M. (2021). Against unitary theories of music evolution. *Behavioral and Brain Sciences*, *44*, e76. https://doi.org/10.1017/S0140525X20001314

Helmholtz, H. L. F. (1885). *On the Sensations of Tone as a Physiological Basis for the Theory of Music* (A. J. Ellis, Trans.). Longmans, Green.

Huang, P., Wilson, M., Mayfield-Jones, D., Coneva, V., Frank, M., & Chitwood, D. H. (2017). *The evolution of Western tonality: A corpus analysis of 24,000 songs from 190 composers over six centuries* [Preprint]. SocArXiv. https://doi.org/10.31235/osf.io/btshk

Huron, D. (1994). Interval-Class Content in Equally Tempered Pitch-Class Sets: Common Scales Exhibit Optimum Tonal Consonance. *Music Perception*, *11*(3), 289–305. https://doi.org/10.2307/40285624

Jordania, J. (2011). *Why do people sing?: Music in human evolution*. Logos: Tbilisi Ivane Javakhishvili State University.

Kopiez, R. (2003). Intonation of Harmonic Intervals: Adaptability of Expert Musicians to Equal Temperament and Just Intonation. *Music Perception*, *20*(4), 383–410. https://doi.org/10.1525/mp.2003.20.4.383

Kulkarni, S., David, S. U., Lynn, C. W., & Bassett, D. S. (2024). Information content of note transitions in the music of J. S. Bach. *Physical Review Research*, *6*(1), 013136. https://doi.org/10.1103/PhysRevResearch.6.013136

Kunst, J. (1967). *Music in New Guinea: Three Studies*. Springer Netherlands. https://doi.org/10.1007/978-94-017-6083-6

Lahdelma, I., & Eerola, T. (2023). Data-driven theory formulation or theory-driven data interpretation?: Comment on "Consonance and dissonance perception. A critical review of the historical sources, multidisciplinary findings, and main hypotheses" by Di Stefano et al. *Physics of Life Reviews*, *45*, 56–59. https://doi.org/10.1016/j.plrev.2023.04.005

Lambert, B., Kontonatsios, G., Mauch, M., Kokkoris, T., Jockers, M., Ananiadou, S., & Leroi, A. M. (2020). The pace of modern culture. *Nature Human Behaviour*, *4*(4), 352–360. https://doi.org/10.1038/s41562-019-0802-4

Lamers, W. S., Boyack, K., Larivière, V., Sugimoto, C. R., Van Eck, N. J., Waltman, L., & Murray, D. (2021). Investigating disagreement in the scientific literature. *eLife*, *10*, e72737. https://doi.org/10.7554/eLife.72737.sa2

Leroi, A. M., Lambert, B., Mauch, M., Papadopoulou, M., Ananiadou, S., Lindberg, S. I., & Lindenfors, P. (2020). On revolutions. *Palgrave Communications*, *6*(1), 1–11. https://doi.org/10.1057/s41599-019-0371-1

Long, B., Simson, J., Buxó-Lugo, A., Watson, D. G., & Mehr, S. A. (2023). How games can make behavioural science better. *Nature*, *613*(7944), 433–436. https://doi.org/10.1038/d41586-023-00065-6

Mann, A. (2011). High-temperature superconductivity at 25: Still in suspense. *Nature*, *475*(7356), 280–282. https://doi.org/10.1038/475280a

Marage, P., & Wallenborn, G. (1999). The Debate between Einstein and Bohr, or How to Interpret Quantum Mechanics. In P. Marage & G. Wallenborn (Eds.), *The Solvay Councils and the Birth of Modern Physics* (pp. 161–174). Birkhäuser Basel. https://doi.org/10.1007/978-3-0348-7703-9_10

Margulis, E. H. (2021). Pluralism provides the best chance for addressing big questions about music. *Behavioral and Brain Sciences*, *44*, e82. https://doi.org/10.1017/S0140525X20000965

Marjieh, R., Harrison, P. M. C., Lee, H., Deligiannaki, F., & Jacoby, N. (2024). Timbral effects on consonance disentangle psychoacoustic mechanisms and suggest perceptual origins for musical scales. *Nature Communications*, *15*(1), 1482. https://doi.org/10.1038/s41467-024-45812-z

McBride, J. M., & Avendaño, C. (2017). Phase behaviour and gravity-directed self assembly of hard convex spherical caps. *Soft Matter*, *13*(10), 2085–2098. https://doi.org/10.1039/C6SM02678H

McBride, J. M., & Avendaño, C. (2018). Plastic fluctuations in empty crystals formed by cubic wireframe particles. *Physical Review Materials*, *2*(5), 055601. https://doi.org/10.1103/PhysRevMaterials.2.055601

McBride, J. M., Passmore, S., & Tlusty, T. (2023). Convergent evolution in a large cross-cultural database of musical scales. *PLOS ONE*, *18*(12), e0284851. https://doi.org/10.1371/journal.pone.0284851

McBride, J. M., & Tlusty, T. (2020). *Cross-cultural data shows musical scales evolved to maximise imperfect fifths* (arXiv:1906.06171). arXiv. https://doi.org/10.31234/osf.io/a97x4

McDermott, J. H., Keebler, M. V., Micheyl, C., & Oxenham, A. J. (2010). Musical intervals and relative pitch: Frequency resolution, not interval resolution, is special. *The Journal of the Acoustical Society of America*, *128*(4), 1943–1951. https://doi.org/10.1121/1.3478785

McDermott, J. H., Schultz, A. F., Undurraga, E. A., & Godoy, R. A. (2016). Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature*, *535*(7613), 547–550. https://doi.org/10.1038/nature18635

Mehr, S. A., Krasnow, M. M., Bryant, G. A., & Hagen, E. H. (2021). Origins of music in credible signaling. *Behavioral and Brain Sciences*, *44*, e60. https://doi.org/10.1017/S0140525X20000345

Milne, A. J., Smit, E. A., Sarvasy, H. S., & Dean, R. T. (2023). Evidence for a universal association of auditory roughness with musical stability. *PLOS ONE*, *18*(9), e0291642. https://doi.org/10.1371/journal.pone.0291642

Moore, B. C. J. (1973). Frequency difference limens for short-duration tones. *The Journal of the Acoustical Society of America*, *54*(3), 610–619. https://doi.org/10.1121/1.1913640

Moss, F. C., Neuwirth, M., & Rohrmeier, M. (2023). The line of fifths and the co-evolution of tonal pitch-classes. *Journal of Mathematics and Music*, *17*(2), 173–197. https://doi.org/10.1080/17459737.2022.2044927

Nakamura, E., & Kaneko, K. (2019). Statistical Evolutionary Laws in Music Styles. *Scientific Reports*, *9*(1), 15993. https://doi.org/10.1038/s41598-019-52380-6

Ozaki, Y., et al. (2024). Globally, songs and instrumental melodies are slower and higher and use more stable pitches than speech: A Registered Report. *Science Advances,* 10, eadmn9797. https://doi.org/10.1126/sciadv.adm9797

Pearson, L. (2016). Coarticulation and Gesture: An Analysis of Melodic Movement in South Indian Raga Performance: Coarticulation and Gesture. *Music Analysis*, *35*(3), 280–313. https://doi.org/10.1111/musa.12071

Pelofi, C., & Farbood, M. M. (2021). Asymmetry in scales enhances learning of new musical structures. *Proceedings of the National Academy of Sciences*, *118*(31), e2014725118. https://doi.org/10.1073/pnas.2014725118

Pfordresher, P. Q. (2022). Singing accuracy across the lifespan. *Annals of the New York Academy of Sciences*, *1515*(1), 120–128. https://doi.org/10.1111/nyas.14815

Pfordresher, P. Q., & Brown, S. (2017). Vocal mistuning reveals the origin of musical scales. *Journal of Cognitive Psychology*, *29*(1), 35–52. https://doi.org/10.1080/20445911.2015.1132024

Phillips, E., & Brown, S. (2022a). Beyond harmonicity: Toward a vocal-melody theory of the origin of musical scales. In J. L. Friedmann (Ed.), *Music in the human experience: Perspectives on a musical species*. Cambridge Scholars Publishing.

Phillips, E., & Brown, S. (2022b). Vocal imprecision as a universal constraint on the structure of musical scales. *Scientific Reports*, *12*(1), 19820. https://doi.org/10.1038/s41598-022-24035-6

Powers, H. S. (2001). Mode. In S. Sadie & J. Tryrell (Eds.), *The New Grove Dictionary of Music and Musicians* (2nd ed.). Macmillan Publishers. https://doi.org/10.1093/gmo/9781561592630.article.43718

Rodgers, P. (1998). Superconductivity debate gets ugly. *Physics World*. https://doi.org/10.1088/2058-7058/11/10/15

Rosenzweig, S., Scherbaum, F., Shugliashvili, D., Arifi-Müller, V., & Müller, M. (2020). Erkomaishvili Dataset: A curated corpus of traditional Georgian vocal music for computational musicology. *Transactions of the International Society for Music Information Retrieval*, *3*(1). https://doi.org/10.5334/tismir.44

Ross, B., & Knight, S. (2019). Reports of equitonic scale systems in African musical traditions and their implications for cognitive models of pitch organization. *Musicae Scientiae*, *23*(4), 387–402. https://doi.org/10.1177/1029864917736105

Sadqi, M., Fushman, D., & Muñoz, V. (2006). Atom-by-atom analysis of global downhill protein folding. *Nature*, *442*(7100), 317–321. https://doi.org/10.1038/nature04859

Sadqi, M., Fushman, D., & Muñoz, V. (2007). Analysis of "downhill" protein folding; Analysis of protein-folding cooperativity (Reply). *Nature*, *445*(7129), E17–E18. https://doi.org/10.1038/nature05645

Sanbonmatsu, D. M., Cooley, E. H., & Butner, J. E. (2021). The Impact of Complexity on Methods and Findings in Psychological Science. *Frontiers in Psychology*, *11*, 580111. https://doi.org/10.3389/fpsyg.2020.580111

Savage, P. E., & Currie, T. E. (2017). *Music perception: No strong evidence to reject innate biases*. https://doi.org/10.31235/osf.io/964ut

Savage, P. E., Loui, P., Tarr, B., Schachner, A., Glowacki, L., Mithen, S., & Fitch, W. T. (2021). Music as a coevolved system for social bonding. *Behavioral and Brain Sciences*, *44*, e59. https://doi.org/10.1017/S0140525X20000333

Savage, P. E., Passmore, S., Chiba, G., Currie, T. E., Suzuki, H., & Atkinson, Q. D. (2022). Sequence alignment of folk song melodies reveals cross-cultural regularities of musical evolution. *Current Biology*, *32*(6), 1395-1402.e8. https://doi.org/10.1016/j.cub.2022.01.039

Sethares, W. A. (1998). *Tuning, Timbre, Spectrum, Scale* (2nd ed). Physica-Verlag NetLibrary, Inc. [distributor]. https://doi.org/10.1007/978-1-4471-4177-8

Shi, R., & Tanaka, H. (2020). The anomalies and criticality of liquid water. *Proceedings of the National Academy of Sciences*, *117*(43), 26591–26599. https://doi.org/10.1073/pnas.2008426117

Simonton, D. K. (2015). Psychology as a Science within Comte's Hypothesized Hierarchy: Empirical Investigations and Conceptual Implications. *Review of General Psychology*, *19*(3), 334–344. https://doi.org/10.1037/gpr0000039

Smit, E. A., & Milne, A. J. (2021). The Need for Composite Models of Music Perception. Consonance in Tuning Systems (Familiar or Unfamiliar) Cannot Be Explained By a Single Predictor. *Music Perception*, *38*(3), 335–336. https://doi.org/10.1525/mp.2021.38.3.335

Spiller, H., Clendinning, E. A., & Spiller, H. (2022). *Focus: Gamelan music of Indonesia* (Third edition). Routledge. https://doi.org/10.4324/9781003014836

Surjodiningrat, W., Sudarjana, P. J., & Susanto, A. (1972). *Tone Measurements of Outstanding Javanese Gamelans in Jogjakarta and Surakarta*. Gadjah Mada University Press.

Ternström, S., & Sundberg, J. (1988). Intonation precision of choir singers. *The Journal of the Acoustical Society of America*, *84*(1), 59–69. https://doi.org/10.1121/1.396875

Trehub, S. (1999). Human Processing Predispositions and Musical Universals. In S. Brown, B. Merker, & C. Wallin (Eds.), *The Origins of Music*. The MIT Press. https://doi.org/10.7551/mitpress/5190.003.0030

Trehub, S. E., Schellenberg, E. G., & Kamenetsky, S. B. (1999). Infants' and adults' perception of scale structure. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(4), 965–975. https://doi.org/10.1037//0096-1523.25.4.965

Verosky, N. J. (2017). Hierarchizability as a Predictor of Scale Candidacy. *Music Perception*, *34*(5), 515–530. https://doi.org/10.1525/mp.2017.34.5.515

Wallace, M. L., Larivière, V., & Gingras, Y. (2012). A Small World of Citations? The Influence of Collaboration Networks on Citation Practices. *PLoS ONE*, *7*(3), e33339. https://doi.org/10.1371/journal.pone.0033339

Weaver, W. (1948). Science and Complexity. *American Scientist*, *36*(4), 536–544. https://doi.org/10.14219/jada.archive.1948.0155

Weisser, S., & Demolin, D. (2013, June 6). Variability as a key concept: When different is the same (and vice versa). *Proceedings of the Third International Workshop on Folk Music Analysis*. Third International Workshop on Folk Music Analysis, Amsterdam, Netherlands.

Youngblood, M. (2019). Conformity bias in the cultural transmission of music sampling traditions. *Royal Society Open Science*, *6*(9), 191149. https://doi.org/10.1098/rsos.191149

Youngblood, M., Ozaki, Y., & Savage, P. E. (2023). Cultural Evolution and Music. In J. J. Tehrani, J. Kendal, & R. Kendal (Eds.), *The Oxford Handbook of Cultural Evolution* (1st ed., p. C42S1-C42N14). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198869252.013.42

Zali, Z., Rein, T., Krüger, F., Ohrnberger, M., & Scherbaum, F. (2023). Ocean bottom seismometer (OBS) noise reduction from horizontal and vertical components using harmonic–percussive separation algorithms. *Solid Earth*, *14*(2), 181–195. https://doi.org/10.5194/se-14-181-2023

Zarate, J. M., Ritson, C. R., & Poeppel, D. (2012). Pitch-interval discrimination and musical expertise: Is the semitone a perceptual boundary? *The Journal of the Acoustical Society of America*, *132*(2), 984–993. https://doi.org/10.1121/1.4733535

Zhou, Z., & Bai, Y. (2007). Analysis of protein-folding cooperativity. *Nature*, *445*(7129), E16–E17. https://doi.org/10.1038/nature05644