For *Oxford Handbook of Human Memory*

(Michael J. Kahana and Anthony D. Wagner, Eds.)

Chapter 11.4 Music and Memory

Daniel J. Levitin

Minerva Schools at Keck Graduate Institute

Lindsay Fleming

McGill University

Abstract.

Music is among a small number of human universals and a strong body of evidence indicates that we remember music far better than other stimuli. This chapter will review what we know about these memory abilities and their scientific underpinnings. The superior memory for many things musical may owe to its status as a highly structured stimulus, and to temporal contiguity effects, including its 1/f or fractal structure that is autocorrelative in nature—as well as the effects of repetition, emotional content, autobiographical salience, and how highly structured the stimulus is. Musical memory functions in many ways like memory for words, pictures, and other experiences. It is consistent with standard models of memory, and with schema theory. The musical memories of professional and non-professional musicians alike are often more detailed, and accurate, and demonstrate a greater capacity than other memory systems. It follows psychological laws of general memory, including levels of processing, serial position effects, and preferential processing for emotionally meaningful stimuli. The primary focus here will be long-term memory, both because so much more is known about it, and because this is the level at which most of us meaningfully engage with music day-to-day via our long-term memory for songs.

Music is among a small number of human universals (Brown, 2000; Nettl, 1983). No culture now nor anytime in the past has lacked music. While there remain ongoing debates about whether music constitutes an evolutionary adaptation, or an evolutionary by-product,[1] there is strong evidence that humans remember music far better than other stimuli. This chapter will review what we know about these memory abilities and their scientific underpinnings.

A number of factors influence the memorability of any stimulus. These include repetition, emotional content, autobiographical salience, and how highly structured the stimulus is. These factors hold for musical memory as well. For example (regarding structure), melodies written by great composers are better remembered than random sequences of notes, and tonal music is better remembered than atonal music (Schulze, Dowling, & Tillmann, 2012). Musical memory fits into the modal memory model (Atkinson & Shiffrin, 1968, 1971) and Baddeley's (1990) update of that model, and can be elaborated with the addition of a musical memory loop added in parallel to the phonological loop (Berz, 1995; it has been shown that there is not simply an "auditory" loop; see Schulze & Koelsch, 2012). This update accounts for the unattended music effect, that unattended instrumental music does not cause disruptions in verbal performance, and Baddeley's own recognition that "the fact we can hear and remember sounds that are very unlike speech means that there must be some additional form of

---

[1] Stephen Jay Gould used the term *spandrel* for evolutionary byproducts, a phenotypic trait that is not a direct product of adaptive selection. Others have claimed that music is an exaptation (a co-opted phenotypic trait that evolved for a purpose other than its current use, such as bird feathers, which were an adaptation for warmth that became co-opted for flight). See, for example, the review by Buss, D. M., Haselton, M. G., Shackelford, T. K., Bleske, A. L., & Wakefield, J. C. (1998). Adaptations, exaptations, and spandrels. *American psychologist*, *53*(5), 533.

acoustic storage system capable of dealing with such material" (Salamé & Baddeley, 1989, p. 121).

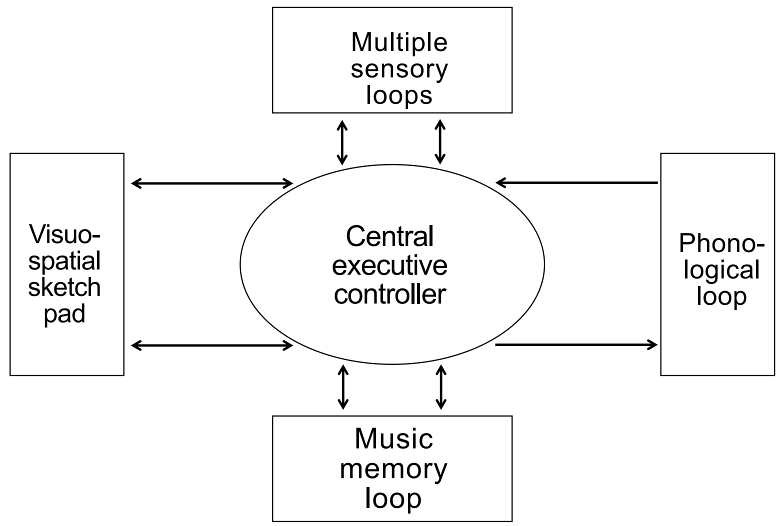

Figure 1. The working memory model of Baddeley (1990) updated by Berz (1995) with a Music Memory Loop.

Musical memory also follows that collection of psychological laws that cognitive psychologists have amassed about memory in general (See Laws of Memory, Chapter 1.2), including levels of processing (Craik & Lockhart, 1972; for music see Segalowitz, et al., 2001), serial position effects (see Serial Recall, Chapter 5.4; Hursltone & Hitch, 2015; for music see Greene & Samuel, 1986; Palmer & Pfordresher, 2003) and preferential processing for emotionally meaningful stimuli (Hu et al, 2007; Tyng et al, 2017; for music see Eschrich, Münte, & Altenmüller, 2008).

The primary focus of this article will be long-term memory, both because so much more is known about it, and because this is the level at which most of us meaningfully engage with music day-to-day via our long-term memory for songs.[2]

The chapter begins with a brief overview and definition of what we mean by "music." The structure and constraints that facilitate musical memory are examined, as are the universals that characterize nearly all music. Real world applications, resolution and durability of memory for music are discussed. The storage and retrieval processes and memory cues for music are then described.

What is music?

Defining music has been difficult due to the diversity of forms of musical expression that are found across the world (Nettl, 1983). Westerners tend to think of music as involving combinations of pitches, but this ignores the prominence and form-bearing importance of rhythm and timbre of many African musics (Chernoff, 1979), and of contemporary hip-hop and funk (Greenwald, 2002; Jowers, 1999). In studying a human universal, it seems prudent to cast a wide net of inclusiveness. As philosopher Stephen Schoeder notes, "The point is not to build and defend a specialized realm so much as to open human experience (and the world in which it happens) to critical and appreciative reflection—to 'enlarge the universe of human discourse' " (c.f. Geertz, 1973). The diversity of forms that music takes in the world is great, from the

[2] For examples of work on musical echoic memory see, Nees (2016); for musical short-term memory, see Aizenman, Gold & Sekuler (2018); Nees, Corrini, Leong & Harris (2017). Music's engagement with the "rehearsal loop" of the modal model is described by Godoy & Jorgensen (2012); Kaernbach & Schlemmer (2008); Nees, et al (2017).

asynchronous and atonal (to Western ears) singing of Cameroon pygmies, to the microtonal and drone-based musics of India, to Tuvan throat-singing and the "cosmic jazz" of Sun Ra in 1960s America, and the taped experiments of Karlheinz Stockhausen, Pierre Schaeffer and Anette Vande Gorgne. And yet, we humans recognize in other culture's musics a kinship to our own. Although we may not understand or even like the music of another culture, we still recognize it as *music*, much as we can recognize the variety of sounds used in disparate languages as human speech. One recent study suggests that we recognize not only the form, but also the function of songs from other cultures (Mehr, Singh, York, Glowacki, & Krasnow, 2018, following the hypothesis first proposed in Levitin, 2008).

Perhaps the most satisfying and inclusive definition of music comes from the composer Edgard Varèse, who defined music as "organized sound." This implies an intentionality — music is not simply a collection of random noises, but something structured (West, Howell & Cross, 1985), and that structure is presumably borne of constraints, both neurocognitive and physical.

Structure and constraints facilitating musical memory

Our perception of music, and thus any subsequent memory for it, is necessarily constrained by innate principles of organization (Kubovy & Pomerantz, 1981/2019; McDermott & Hauser, 2005) and by multiple layers of schematic representations (Bharucha, 1987). Through a lifetime of listening, we acquire highly developed schemas for music, pitch relations, chords, scales, rhythms, tempos, loudness changes, harmony, phrase structure, and musical styles. These musical schemas resemble scripts in verbal storytelling (Rumelhart, 1975; Schank & Abelson, 1977). We know, for example, that a "blues" will have a recognizable chord structure (of I^7, IV^7, V^7), it will have a 12-measure or 16-measure phrase structure, and the lyric of the first line will tend to repeat in the second line of the verse ("I went down to the crossroads — fell down on my knees/I went down to the crossroads — fell down on my knees…"). These forms may be known by most members of a culture only implicitly, and lacking formal descriptors, but they are still known in the Chomskian sense of us all implicitly knowing the grammar and standard structures for our native language and dialect and being able to detect departures from those stylistic norms (Chomsky, 1967). Paralleling Chomsky's innate generative syntactic capacity, generative processes in music have been proposed (Lerdahl & Jackendoff, 1996).

Among the structural universals found across all the world's musics are the octave (a frequency ratio of 2:1), and the perception that two notes that are an octave apart are, in an important sense, the *same note* (Dowling & Harwood, 1986). Indeed, when we say that men and women are singing or speaking in unison, their voices are typically an octave apart, and so technically, *not* in unison, but we regard them as equivalent. This mental perception is informed by properties of the natural world: two tones an octave apart have identical overtone structures except for one element. Secondly, most cultures also have the Perfect Fifth, an interval ratio of 3:2, a reflection of the physical acoustics of the overtone series generated by vibrating objects. Thirdly, most cultures divide the octave up into a discrete number of tones, or scale degrees, typically 5 – 9, reflecting closely the capacity limitations put forth by Miller (1956). These factors provide constraints on the design of musical instruments, on the creation of musical compositions and their performance, and then on subsequent learning and recall.

Among those cultures that use more than 9 notes, such as the Western diatonic scale of 11, usually only 7 notes are used at any one time, or in any one passage. The relations between notes of a scale provide additional constraints such that most casual listeners familiar with their culture's own music can make accurate predictions about what notes will come next in a song they've never heard before (Krumhansl & Kessler, 1982) and even toddlers can tell when a note or chord is out of place (Koelsch, et al., 2005). Additional organizational principles of music are believed to reflect underlying, universal principles of cognition, including a modality-general time-distance law, alongside latent cognitive structures that correspond to geometrically regular configurations in higher-dimensional spaces (Janata, et al, 2002; Shepard, 1995; 2009; Tymoczko, 2006).

The highly structured nature of music facilitates schema abstraction, which in turn facilitates both storage and retrieval processes. As Preston & Varga (Chapter 6.6, this volume) note, "schemas extract goal-relevant features that are common across individual episodes… schemas consist of context-specific associative structures that are abstracted across multiple events."

Savage, et al. (2015) identified 18 universals that characterize nearly all of the world's music. These necessarily interact with constraints of the human memory system as both music and memory have co-evolved. Those same constraints facilitate storage and recall processes by limiting the forms and types of information that become memories. The 12 of Savage's 18 universals most relevant here are related to pitch, rhythm, and form.

- Pitch. Music generally uses discrete pitches (1) to form nonequidistant scales (2) containing seven or fewer scale degrees per octave. (3) Most music uses descending or arched melodic contours (4) composed of small intervals (5) of a perfect fifth or smaller (roughly half an octave).
- Rhythm. Music tends to (6) use an isochronous beat (7) organized according to metrical hierarchies (8) based on multiples of two or three beats, (9) especially multiples of two beats. (10) This beat tends to be used to construct motivic patterns (11) based on fewer than five durational values.
- Form. (12) Music tends to consist of short phrases less than 9 s long, conforming to the limits of echoic memory with ongoing interference.

Thus, memory for music scaffolds on hierarchical representations for music's structural components, and related mental models and schemata (Clarke, Dibben & Pitts, 2010; Deliége, Mélen, Stammers & Cross, 1996; Howell, West & Cross, 1991; Krumhansl & Castellano, 1983). Such representations are believed to rely on a number of principles found elsewhere in cognition, including Gestalt principles of prägnanz, similarity, closure, good continuation, and common fate (Deutsch, 1982; Dowling & Harwood, 1986; Leman, 1997; Lerdahl & Jackendoff, 1996; Shepard & Levitin, 2002; Terhardt, 1987). In addition, the mutually reinforcing constraints of rhythm, tempo, pitch, contour, melody, tonality, key, rhyme, and accent structure serve to limit the number of possible notes or lyrics that could complete a forgotten phrase, allowing for on-line reconstitution of forgotten parts. These principles and structural constraints can only facilitate memory for music that is heard in styles and tuning systems that one is familiar with (Herff, et al, 2017).

With respect to rhythm, the sorts of hierarchically embedded internal clocks or oscillators that govern timekeeping in everyday tasks (e.g. Breska & Ivry, 2016; Gallistel & Gibbon, 2000) are engaged in musical memory encoding and retrieval tasks, and in musical listening (Brochard, et al, 2003; Janata & Grafton, 2003; Levitin, 2009). Phase-locking and tempo tracking allow musicians to keep track of multiple musical events simultaneously and to recover from errors while staying in time. It's what allows listeners to continue singing along with a song, even when their satellite radio signal goes out under an overpass or they forget a word or two. And it facilitates memory storage and retrieval by governing when important metrical events must occur.

Music performance, listening and attendant memory are subject to temporal order constraints. Following Lashley (1951), temporal sequences and musical rhythms are related to motor function, linking rhythmic action, perception and motor organization. Related constraints apply, such as the dynamics of inertia (Martin, 1972).

Musical elements are not stochastically concatenated, but bear an internal (and largely logical) relation to one another. Accents and barlines (the superordinate level above the beat) provide part of the internal structure, as do phrases. In effect, musical phrases function like sentences in spoken language (Caplin, 2000). Starting and stopping points in musical recall are not arbitrary; musicians and non-musicians alike tend to start and stop at consistent, canonical points within the sequence, and these points are hierarchically organized (Hyman & Rubin, 1990; Rubin, 1995). Errors in recollection of musical sequences thus tend not to be evenly distributed throughout a piece, but to occur at structural boundaries (Chaffin & Imreh, 2002; Deutsch, 1980; Deutsch & Feroe, 1981; Mishra, 2010). The temporal mediation of memory I'm describing here parallels findings in word memory: even for lists of random words, or words from a small number of different categories, people typically prefer to recall items in the order in which they were presented even when they don't have to (Mandler, 1969; Postman, 1972).

A brief mathematical interlude: Fractals and self-similarity in music that reinforce structural memory

One way in which musical structure incorporates physical regularities of the natural world is with its autocorrelative properties, modeled by a 1/f power law (otherwise known as "fractal"), similar to that governing a wide range of natural phenomena including the topography of coastlines, neural spike patterns, and the structure of fern fronds (e.g. Mandelbrot, 1983). Fractal structures are characterized by self-similarity; at different levels of resolution, they resemble themselves. Thus, a single frond from a fern may have a triangular shape. On closer inspection, one sees that the frond is made up of triangularly-shaped clusters (leaflets, or *pinnae)*, and those clusters are made up of triangularly-shaped components (subleaflets, or *pinnule),* and that those components in turn are made up of triangularly-shaped lobes.

Figure 2. Lady Fern (Athyrium filix-femina) at Muir Woods, California.

Music is known to embody such self-similar properties across time. The fractal nature of music, and the corresponding 1/f power law, has been documented using mathematical analyses in the domains of pitch (Voss & Clarke, 1975), rhythm (Levitin, Chordia & Menon, 2012) and harmony (Wu, et al, 2015). This self-similarity is a latent property that listeners are typically unaware of.

Self-similar systems are recursive, in that they are defined by reference to themselves, a concept popularized in computer science when a defined function or subroutine calls upon itself. Human cognition shows recursive hierarchies in a variety of domains, including linguistic, visual perception, spatial perception, music, and memory. This appears to be an ability that emerges in infancy around 12–14 months (Lewkowicz, Schmuckler & Mangalindan, 2018). Memory in general (including musical memory) is recursive insofar as both storage and retrieval depend on context (Anderson, Pichert & Shirey, 1983; Bransford & Franks, 1971). Wachter and Kahana (2019) modeled this property for human memory, noting that retrieval cueing with context allows us to recover the items associated with that context, and cueing with an item (a feature) recovers the contexts previously associated with that item, and thus demonstrates recursion.

Consider a typical, everyday instance of recursion while attempting to remember an individual instance which forms a set of repeated instances (a walk in the park, ordering salad niçoise at your favorite restaurant). If one's encounter with an item includes being reminded of the previous encounter, and if this compound experience is encoded into memory, then on a future third encounter, one remembers having remembered (or is reminded of having been reminded; Hintzman 2011). Thus recursion, in which remindings are embedded in remindings, are built up over multiple representations. (See also e.g., Negley, Kelley & Jacoby, 2018, for laboratory studies of recursive memory with word lists).

Real world memory for music

Travelling bards over the centuries traditionally lacked literacy, and yet sang long epic poems from memory; in many parts of the world they still do. Some of these epic songs are thousands of lines long and take hours to sing (*The Iliad* and *The Odyssey,* for example, total 27,000 lines). On the surface, then, these would appear to be extraordinary demonstrations of memory. Yet in pre-literate and illiterate cultures, the notion of verbatim memory is different than what we hold in the experimental psychology laboratory— the very idea of a verbatim (literal) remembering is a product of literacy itself (Lord, 1960) and was not the goal of those telling epic stories through song. Neisser (1982) concluded that none of what appear to be unusual feats of lyrical memory fits the stereotype of "rote memory" studied in modern laboratories at all: epic poems are reinvented each time they are sung.

Today, verbatim lyrical memory is typically expected. Whether we're singing along with "The Star Spangled Banner" or "Single Ladies (Put A Ring On It)", we know that we're expected to sing exactly those words and no others, and generally we do. Few of us sat down and studied these lyrics by rote (even Beyoncé uses a teleprompter). We don't rehearse them before we go to a ball game or a pub. The mutually reinforcing constraints mentioned above facilitate storage in a naturalistic fashion. And, errors in musical lyric recall tend to preserve the rhythm, the local and gist meaning, and the poetic sound pattern of the correct lyrics. For example, in a study of everyday recall of Beatles tunes (Hyman & Rubin, 1990), individuals unselected for musical ability made these perfectly reasonable substitutions:

And you know you *should be glad* -- > And you know you *can't be sad*

If you say you love me *too* -- > If you say you love me *true*

To help with good Rocky's *revival* -- > To help with good Rocky's *survival*

<u>Number of songs that can be recalled</u>

Among musicians who play "casuals" — weddings, bar mitzvahs, parties — as well as those who play in hotel lounges and piano bars across the country, it is rather the norm, not the exception, to be able to play thousands of songs from memory. Indeed, this is widely seen as a job requirement. In this respect, the memory of such musicians seems exceptional, standing alongside professional chess players who can remember thousands of different games, or comedians who can remember thousands of jokes. But what's interesting is that we rarely encounter novelists who can recite, say, *Moby Dick* verbatim, or poets who can recite all the works of John Donne. There are no known reported cases of painters or sculptors who can reproduce from memory the works of the great masters. [A complete discussion of exceptional memory is found in Chapter 8.6 by Hambrick & Campitelli in the current volume.]

Musicians who do play popular songs from memory are likely not playing them with true fidelity to a recording or live performance they heard. Rather, they've extracted the gist, higher order features such as style, tonality, and tempo, and they improvise notes that are reasonable and consistent with this higher order schema. They may frequently make chord substitutions that remain syntactically correct for the ones they heard in the original version, not because they can, but because remembering such details rarely detracts from the recognizability of a well-known song. Their errors tend to be structure-preserving, and that memory encoding relies on schematic tonal and rhythmic structures and relations (Sloboda, Hermelin, O'Connor, 1985). In this respect, they resemble the many studies demonstrating the schematic and gist nature of memory for stories and texts (Anderson, Pichert & Shirey, 1983; Bransford & Franks, 1971; Neisser, 1982).

Consider a hypothetical performance by a cover band (a band that plays popular hit songs) in a bar on a Friday night. In a 35-minute set, the musicians may well play 10,000 notes (Clarke, Dibben & Pitts, 2010), but they need not "remember" nearly that many. First, music involves redundancy; once a musician has learned the chorus of the song, s/he doesn't need to memorize each repeat of it, just how many repeats there are. Second, musicians tend to remember chord sequences, rather than individual chords (Dowling, 1973), akin to textual or numeric chunking (Gobet et al, 2001; Mewhort, 1972), allowing for a relatively sparse schematic representation of each song (Clarke, Dibben & Pitts, 2010).

Performance expectations for classical musicians, particularly soloists, are different than those for popular music performers, and classical musicians are indeed expected to play the music *as written*. But this still does not require that they necessarily memorize each individual note — they employ the same strategies mentioned above.

I interviewed three Grammy-winning musicians for this chapter. Composer and arranger Chris Walden remembers having to turn in a list of 100 memorized jazz standards as part of his musical training at the Conservatory of Cologne. I asked three-time Grammy-winning pianist Shelly Berg — who by the way, is extremely self-effacing and always errs on the side of modesty — how many songs he thought he could play from memory. He answered without hesitation, "2500." And how many more could he remember well enough to fake his way through them so that most people wouldn't know the difference? "Another 2500," he said. World-renowned harpist and violinist Carlos Reyes, overhearing the conversation added that "I find I'm more likely to be able to remember a song in a band situation than if I'm playing solo. The drummer might start a groove and then I realize, 'Oh yeah, I know that song." Shelly Berg added, "If the bass player knows the song all I have to do is listen to him and then I know the song." (Walden, Berg, Reyes, personal communication, 2019). Although these are only anecdotal, they are the kinds of observations out of which scientific experiments grow. More work needs to be formally conducted on this topic.

Musicians and non-musicians alike often describe such demonstrations of fluid musical production as "motor memory" or "muscle memory." The latter is a misnomer because the peripheral muscles themselves don't remember the movements, but the motor system as a whole does — a system that involves motor action planning, sequence construction and binding, and motor movements, and entails pre-motor cortex, supplementary motor cortex, motor cortex, sensory cortex, the hippocampus, the basal ganglia and the cerebellum, among other regions. And yet it can't strictly be motor memory that allows for prodigious musical production memory, because many multi-instrumentalists can play a piece on an instrument they've never played it on before (Clarke, Dibben & Pitts, 2010). This suggests that the memory representation may exist at two levels of abstraction — a higher level that encodes auditory memory and a description of intervals and timing, and a subordinate level that encodes the specific movements on a specific instrument to achieve those. In both cases, an internal representation or mental map must exist that maps specific tones of the scale to specific movements of the fingers (or feet, or vocal cords), or that maps tones to specific locations on an instrument.

Resolution and durability

Halpern (1989) asked nonmusicians to sing well-known songs such as “Happy Birthday” or “Frère Jacques” from memory on two different occasions. She found that although people tended not to sing in the same keys as one another, they did tend to sing a song consistently, in the same key from one occasion to another. This suggested that they had encoded the pitches of the songs in long-term memory. Perhaps these results could be accounted for without invoking pitch memory if the participants had simply relied on "muscle memory" for the tension in their vocal cord from one testing session to another. Of course, muscle memory — or to be precise, motor cortex memory — is still a form of memory and labeling the phenomenon does nothing to change it. But Ward and Burns (1978) showed that muscle memory isn’t actually very good; when deprived of auditory feedback, skilled singers could only get to within a third of an octave of the correct tone. Cook (1996) showed that the same was true for skilled trombonists who were denied auditory feedback and had to rely on "muscle memory".

The songs used in Halpern's study lack a standardized canonical key in which to be performed. A follow-up study adapted Halpern's paradigm to songs that exist in primarily only a single version, that may be heard hundreds or thousands of times, contemporary (at the time of the study, in 1990) popular and hit songs such as "New Years Day" by U2 or "Material Girl" by Madonna. Participants, unselected for musical ability, tended to sing their selected songs at or very near the correct pitch (Levitin, 1994; see Frieler, et al., 2013 for a replication).

Using the same data set, Levitin & Cook (1996) showed that the participants had reproduced the tempo of those songs with great accuracy, with a Pearson's correlation of 0.95. On Trial 1, 72% of the participants performed within ±4% of the actual tempo for the songs, yielding a D' of 1.6. In prior work, the JND for tempo has been found to be around 6%, putting these responses within less than one JND. 88% of participants performed with ±8%, yielding a D' of 2.4. These results may even underestimate the strength of tempo memory, because subjects were only instructed to reproduce pitches accurately; to the extent that they also reproduced tempo, they did this on their own, without being requested by the experimenter to do so. In listening to the recordings of the participants overlayed with the recordings of the songs they were attempting to sing, one is struck by how closely the two line up. It is as though they are singing along with the recording and in effect they are: they are singing along with a memory of the song in their mind's ear that displays the astonishing fidelity of those memories.

Subsequent studies have shown that long-term musical memory can retain highly detailed spectral and timbral information, such as memory for loudness (Tse & Levitin, 2019), memory for timbre when melody and rhythm are removed (Krumhansl, 2010; Schellenberg, Iverson & Mckinnon, 1999), or the sound of a song being filtered at 1.2 KHz by 2 dB, within the level of the JND for spectrum perception (Quesnel, 1996; 2002).

Much work on musical memory in the first one hundred years of experimental psychology (roughly from 1870 to 1970) was hobbled by researchers' use of artificial stimuli, often lacking ecological validity or even qualities that we would consider "musical." In an effort to control the acoustical characteristics of stimuli, researchers often composed their own melodies, and if those were not easily recalled by experimental participants, it was interpreted as a failure of human memory, not for what it seems self-evident it was: a failure of a non-professional composer to write a memorable melody (Tirovolas & Levitin, 2011).

Storage and retrieval processes; memory cues

Christian von Ehrenfels (1890) posed a question that launched the Gestalt psychology movement: How is it that we can change all of the notes in a melody — such as under transposition — and yet still recognize that melody? We must be extracting the temporal-spectral patterns of music and forming (and possibly storing) an abstract representation of them. Adaptive behavior requires stimulus generalization, and distinguishing those properties of a stimulus that are invariant and essential to its identity from those that are situationally variant, reflecting momentary differences in presentation (Shepard, 2009; Tooby & Cosmides, 2007). Song recognition requires that we ignore some stimulus features while we focus only on higher order features that are invariant from one listening to the next—and in this way, extract invariant properties of a song. Changes in volume, spatial location, and listening context (such as whether the sound is coming from live instruments, headphones or speakers) do not affect the identity of a song; although the evidence reviewed in the previous section shows that these features may become encoded in memory, they are not the primary features required for identification and recognition memory of music.

White (1960) took up this question by creating a number of systematic transformations of well-known songs, in time and in pitch, and found that even quite extreme departures from the canonical version of a song did not interfere with its recognizability. Musical memory thus appears to be made possible by the schema abstraction first studied by Posner & Keele (1968, 1970). Consistent with this, and in a specifically musical context, false positive judgments increase with similarity to a melodic prototype previously presented (Deliége, 2001).

Retrieval cues

The multi-dimensional nature of music allows for a variety of retrieval cues to be effective for song identification. The once-popular radio and television show "Name That Tune" was like an ongoing psychological laboratory, in which contestants, often astonishingly, could name a song on the basis of only a few notes — sometimes just one note. A representative clue was:

> "Contrary to the title, Perry Como had no difficulty earning a 1970s gold disc with this emotional tune."

In this episode from 1978, a contestant was able to correctly name *It's Impossible* with just five notes (B, A, C, B, B; Edwards, 1978).

Lyrics, semantic themes, pitch intervals, rhythms, timbre[3] and chords are just a few of the sorts of retrieval cues that function to bring forth vivid memories of songs. When participants are asked to name as many songs as they can on a particular semantic theme — say "cars" or "divorce" — they can generate dozens on the spot (Levitin, 2005). Certain melodic sequences are sufficiently unique that they can provoke immediate

[3] Timbre is the perceived sound quality of a musical sound or tone. Timbre is what distinguishes two instruments or voices from one another when they are playing the same note at the same volume for the same duration. It is itself a multi-dimensional construct formed by the auditory system from spectro-temporal qualities, including pitch information in the overtone series, loudness, and changes in those qualities over time. Various models have been proposed, including a three dimensional model that describes acoustic properties (*attack time, spectral centroid,* and *spectrum fine structure;* Caclin, et al, 2005) and a five-dimensional model describing perceptual qualities (*hard/soft, sharp/dull, high-/low-frequency, energy balance, explosive/calm;* Elliot, Hamilton & Theunissen, 2013).

identification. The first six notes of *Over The Rainbow*, the first three notes of *Yesterday,* one measure of the rhythm of *The William Tell Overture* or of *De Camptown Races* are high in cue validity and cue item discriminability (Rosch, 1999) and are not easily confusable with other songs (Houlihan & Levitin, 2011).

Even timbre alone, in the absence of melody or rhythm, can allow for accurate and rapid song identification. In the spirit of the Gestalt psychologists, most of whose "experiments" were really demonstrations, this author has performed a naturalistic experiment with an *n* of several thousands of students and lecture attendees over a decade. These experiments consist of playing short (200–500 ms) excerpts of songs and the audience is asked to shout out the name of the song. Such excerpts are edited so that melody and rhythm are completely lacking, because the excerpts are shorter in duration than any one note of the song. With songs as diverse as *Eleanor Rigby* (Beatles)*, Bennie and the Jets* (Elton John) and *Don't Know Why* (Nora Jones), the majority of the audience identifies the song within one second of hearing the excerpt. There are age effects as well. When played a 500 ms excerpt of *Work* (Rihanna, feat. Drake), almost all audience members under the age of 27 (as of this writing) can identify it, while few people over 40 can.

A more recent, and formal demonstration used 300 and 400 ms excerpts from popular music spanning five decades, from the 1960s to the 2000s (Krumhansl, 2010). For 400 ms clips, participants identified both artist and title on more than 25% of the trials. Very accurate confidence ratings showed that this knowledge was recalled consciously. Performance was somewhat higher if the clip contained a word or part word from the title. Even when a clip was not identified, it conveyed information about emotional content, style and, to some extent, decade of release. A related study used recognition rather than identification (that is, listeners were given a list of possible songs rather than having to generate the names on their own). Listeners were presented with 200 ms excerpts of well-known popular songs and were able to identify them at levels well above chance (Schellenberg, Iverson & Mckinnon, 1999), and for songs that are known but not necessarily identifiable, a feeling of familiarity can be triggered by 500 ms excerpts (Filipic, Tillman & Bigand, 2010).

Finally, another effective retrieval cue for music is emotion. It is well established that emotional events are more frequently and vividly remembered than neutral ones (see Chapter 3.7 by Kensinger in this volume). Because music is capable of evoking a wide variety of strongly felt emotions (Dowling & Harwood, 1986) this also accounts in part for why older adults with dementia or cognitive impairment are often able to remember songs from their teenage years when other memory systems have failed; the other part of the story is the multiple overlapping structural constraints of rhyme, rhythm, meter, and accent structure (for more an aging and memory see Chapter 8.4 by Light, this volume).

Summary

Musical memory functions in many ways like memory for words, pictures, and other experiences. It is consistent with standard models of memory, and with schema theory. The musical memories of professional and non-professional musicians alike are often more detailed, and accurate, and demonstrate a greater capacity than other memory systems, although such direct comparisons have not been the subject of rigorous study.

The superior memory for many things musical may owe to its status as a highly structured stimulus, and to temporal contiguity effects, including its 1/f or fractal structure that is autocorrelative in nature.

Much of what we have come to know about musical memory comes from naturalistic, observational studies of musicians in real world settings, and in the past several decades, controlled laboratory studies have expanded this knowledge. The majority of these true experiments have focused on long-term memory for music. There is still much work to be done in this domain, and in understanding more about echoic and short-term musical memory.

References

Aizenman, A. M., Gold, J. M., & Sekuler, R. (2018). Multisensory integration in short-term memory: Musicians do rock. *Neuroscience*, *389*, 141-151.

Anderson, R. C., Pichert, J. W., & Shirey, L. L. (1983). Effects of the reader's schema at different points in time. *Journal of Educational Psychology*, *75*(2), 271.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In *Psychology of learning and motivation* (Vol. 2, pp. 89-195). Academic Press.

Atkinson, R., & Shiffrin, R. (1971). The control processes of short-term memory (Technical report 173, April 19, 1971). *Institute for Mathematical Studies in the Social Sciences, Stanford University*.

Baddeley, A. (1990). *Human memory: Theory and practice*. Boston: Allyn and Bacon.

Berz, W. L. (1995). Working memory in music: A theoretical model. *Music Perception: An Interdisciplinary Journal*, *12*(3), 353-364.

Bharucha, J. J. (1987). Music cognition and perceptual facilitation: A connectionist framework. *Music Perception: An Interdisciplinary Journal*, *5*(1), 1-30.

Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. *Cognitive psychology*, *2*(4), 331-350.

Breska, A., & Ivry, R. B. (2016). Taxonomies of timing: where does the cerebellum fit in? *Current opinion in behavioral sciences*, *8*, 282-288.

Brochard, R., Abecasis, D., Potter, D., Ragot, R., & Drake, C. (2003). The "ticktock" of our internal clock: Direct brain evidence of subjective accents in isochronous sequences. *Psychological Science*, *14*(4), 362-366.

Brown, D. E. (2000). Human universals and their implications. In N. Roughley (Ed.) *Being humans: Anthropological universality and particularity in transdisciplinary perspectives*. New York: Walter de Gruyter.

Buss, D. M., Haselton, M. G., Shackelford, T. K., Bleske, A. L., & Wakefield, J. C. (1998). Adaptations, exaptations, and spandrels. *American psychologist*, *53*(5), 533. [n.b. cited in Footnote 1]

Caclin, A., McAdams, S., Smith, B. K., & Winsberg, S. (2005). Acoustic correlates of timbre space dimensions: A confirmatory study using synthetic tones. *The Journal of the Acoustical Society of America*, *118*(1), 471-482.

Caplin, W. E. (2000). *Classical form: A theory of formal functions for the instrumental music of Haydn, Mozart, and Beethoven*. Oxford University Press.

Chaffin, R. and Imreh, G. (2002). Practicing perfection: piano performance as expert memory. *Psychological Science, 13*, 342–349.

Chernoff, J. M. (1979). *African rhythm and African sensibility*(Vol. 36). Chicago: University of Chicago press.

Chomsky, N. (1967). Recent contributions to the theory of innate ideas. In *A Portrait of Twenty-five Years* (pp. 31-40). Springer, Dordrecht.

Clarke, E., Dibben, N., & Pitts, S. (2010). *Music and mind in everyday life*. Oxford University Press.

Cook, P. (1996). Hearing, feeling, and performing: Masking studies with trombone players. In *International Conference on Music Perception and Cognition, Montreal*. https://www.cs.princeton.edu/~prc/HearFeel.pdf

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of verbal learning and verbal behavior*, *11*(6), 671-684.

Deliége, I. (2001). Prototype effects in music listening: An empirical approach to the notion of imprint. *Music Perception: An Interdisciplinary Journal*, *18*(3), 371-407.

Deliége, I., Mélen, M., Stammers, D., & Cross, I. (1996). Musical schemata in real-time listening to a piece of music. *Music Perception: An Interdisciplinary Journal*, *14*(2), 117-159.

Deutsch, D. (1980). The processing of structured and unstructured tonal sequences. *Perception & Psychophysics, 28*, 381–389.

Deutsch, D. (1982). Organizational processes in music. In *Music, mind, and brain* (pp. 119-136). Springer, Boston, MA.

Deutsch, D., & Feroe, J. (1981). The internal representation of pitch sequences in tonal music. *Psychological review*, *88*(6), 503.

Dowling, W. J. (1973). Rhythmic groups and subjective chunks in memory for melodies. *Perception & Psychophysics*, *14*(1), 37-40.

Dowling, W. J., & Harwood, D. L. (1986). *Music cognition*. San Diego: Academic Press.

Edwards, R. (Producer). (1978). *Name That Tune* [Television series]. Los Angeles, CA: KTTV.

Elliott, T. M., Hamilton, L. S., & Theunissen, F. E. (2013). Acoustic structure of the five perceptual dimensions of timbre in orchestral instrument tones. *The Journal of the Acoustical Society of America*, *133*(1), 389-404.

Eschrich, S., Münte, T. F., & Altenmüller, E. O. (2008). Unforgettable film music: the role of emotion in episodic long-term memory for music. *BMC neuroscience*, *9*(1), 48.

Filipic, S., Tillmann, B., & Bigand, E. (2010). Judging familiarity and emotion from very brief musical excerpts. *Psychonomic Bulletin & Review*, *17*(3), 335-341.

Frieler, K., Fischinger, T., Schlemmer, K., Lothwesen, K., Jakubowski, K., & Müllensiefen, D. (2013). Absolute memory for pitch: A comparative replication of Levitin's 1994 study in six European labs. *Musicae Scientiae*, *17*(3), 334-349.

Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological review*, *107*(2), 289.

Geertz, C. (1973). *The interpretation of cultures* (Vol. 5019). Basic books.

Gobet, F., Lane, P. C., Croker, S., Cheng, P. C., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in cognitive sciences*, *5*(6), 236-243.

Godoy, R. I., & Jorgensen, H. (Eds.). (2012). *Musical imagery*. Routledge.

Greene, R. L., & Samuel, A. G. (1986). Recency and suffix effects in serial recall of musical stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(4),

Greenwald, J. (2002). Hip-hop drumming: The rhyme may define, but the groove makes you move. *Black Music Research Journal*, *22*(2), 259-271.

Halpern, A. R. (1989). Memory for the absolute pitch of familiar songs. *Memory & Cognition*, *17*(5), 572-581.

Herff, S. A., Olsen, K. N., Dean, R. T., & Prince, J. (2017). Memory for melodies in unfamiliar tuning systems: Investigating effects of recency and number of intervening items. *The Quarterly Journal of Experimental Psycholog*y. Advance online publication. http://dx.doi.org/10.1080/ 17470218.2017.1333519

Hintzman, D. L. (2011). Research strategy in the study of memory: Fads, fallacies, and the search for the "coordinates of truth". *Perspectives on Psychological Science*, *6*(3), 253-271.

Houlihan, K., & Levitin, D. J. (2011, August). *Recognition of melodies from rhythm and pitch*. Paper presented at the 2011 bi-annual meeting of the Society for Music Perception and Cognition (SMPC), Rochester, NY.

Howell, P., West, R., & Cross, I. (1991). *Representing musical structure*. Academic Press.

Hu, H., Real, E., Takamiya, K., Kang, M. G., Ledoux, J., Huganir, R. L., & Malinow, R. (2007). Emotion enhances learning via norepinephrine regulation of AMPA-receptor trafficking. *Cell*, *131*(1), 160-173.

Hurlstone, M. J., & Hitch, G. J. (2015). How is the serial order of a spatial sequence represented? Insights from transposition latencies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(2), 295–324.

Hyman, I. E., & Rubin, D. C. (1990). Memorabeatlia: A naturalistic study of long-term memory. *Memory & Cognition*, *18*(2), 205-214.

Janata, P., Birk, J. L., Van Horn, J. D., Leman, M., Tillmann, B., & Bharucha, J. J. (2002). The cortical topography of tonal structures underlying Western music. *Science*, *298*(5601), 2167-2170.

Janata, P., & Grafton, S. T. (2003). Swinging in the brain: shared neural substrates for behaviors related to sequencing and music. *Nature neuroscience*, *6*(7), 682.

Jowers, P. (1999). Timeshards: repetition, timbre, and identity in dance music. *Time & Society*, *8*(2-3), 381-396.

Kaernbach C, Schlemmer K. (2008) The decay of pitch memory during rehearsal. *J Acoust Soc Am.* 123(4):1846-9. doi: 10.1121/1.2875365. PMID: 18396991.

Koelsch, S., Fritz, T., Schulze, K., Alsop, D., & Schlaug, G. (2005). Adults and children processing music: An fMRI study. *NeuroImage, 25,* 1068 – 1076. doi:10.1016/j.neuroimage.2004.12.050

Krumhansl, C. L. (2010). Plink: "Thin slices" of music. *Music Perception: An Interdisciplinary Journal*, *27*(5), 337-354.

Krumhansl, C. L., & Castellano, M. A. (1983). Dynamic processes in music perception. *Memory & Cognition*, *11*(4), 325-334.

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, *89*(4), 334.

Kubovy, M. & Pomerantz, J. (1981/2019). *Perceptual Organization.* Routledge.

Lashley, K. S. (1951). *The problem of serial order in behavior* (Vol. 21). Bobbs-Merrill.

Lewkowicz, D. J., Schmuckler, M. A., & Mangalindan, D. M. (2018). Learning of hierarchical serial patterns emerges in infancy. *Developmental psychobiology, 60*(3), 243-255.

Leman, M. (1997). *Music, gestalt, and computing*. Berlin: Springer.

Lerdahl, F., & Jackendoff, R. S. (1996). *A generative theory of tonal music*. MIT press.

Levitin, D. J. (1994). Absolute memory for musical pitch: Evidence from the production of learned melodies. *Perception & Psychophysics*, *56*(4), 414-423.

Levitin, D. J. (2005). *Recommendation engines and the 5 cent solution.* Future of Music Coalition Fifth Annual Policy Summit, George Washington University, Washington, D. C.

Levitin, D. J. (2008). *The World in Six Songs: How The Musical Brain Created Human Nature.* New York: Dutton/Penguin.

Levitin, D. J. (2009). The neural correlates of temporal structure in music. *Music and medicine*, *1*(1), 9-13. Levitin, D. J. (2019). Absolute Pitch. Chapter 9 in P. J. Rentfrow and D. J. Levitin (Eds), *Foundations in Music Psychology: Theory and Research*. MIT Press, p. 357.

Levitin, D. J., Chordia, P., & Menon, V. (2012). Musical rhythm spectra from Bach to Joplin obey a 1/f power law. *Proceedings of the National Academy of Sciences*, *109*(10), 3716-3720.

Levitin, D. J., & Cook, P. R. (1996). Memory for musical tempo: Additional evidence that auditory memory is absolute. *Perception & Psychophysics*, *58*(6), 927-935.

Lord, A. B. (1960). *The singer of tales.* Cambridge, MA: Harvard University Press.

Mandelbrot, B. B. (1983). *The fractal geometry of nature* (Vol. 173, p. 51). New York: WH freeman.

Mandler, G. (1969). Input variables and output strategies in free recall of categorized lists. *The American Journal of Psychology*, *82*(4), 531-539.

Martin, J. G. (1972). Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psychological Review, 79*(6), 487–509.

McDermott, J., & Hauser, M. (2005). The Origins Of Music: Innateness, Uniqueness, And Evolution. *Music Perception, 23*(1), 29-59. doi:10.1525/mp.2005.23.1.29

Mehr, S. A., Singh, M., York, H., Glowacki, L., & Krasnow, M. M. (2018). Form and function in human song. *Current Biology, 28*(3), 356-368.

Mewhort, D. J. (1972). Scanning, chunking, and the familiarity effect in tachistoscopic recognition. *Journal of Experimental Psychology*, *93*(1), 69.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, *63*(2), 81.

Mishra, J. (2010). Effects of structure and serial position on memory errors in musical performance. *Psychology of Music*, *38*(4), 447-461.

Nees, M. A. (2016). Have we forgotten auditory sensory memory? Retention intervals in studies of nonverbal auditory working memory. *Frontiers in psychology*, *7*, 1892.

Nees, M. A., Corrini, E., Leong, P., & Harris, J. (2017). Maintenance of memory for melodies: Articulation or attentional refreshing?. *Psychonomic bulletin & review*, *24*(6), 1964-1970.

Negley, J. H., Kelley, C. M., & Jacoby, L. L. (2018). The importance of time to think back: The role of reminding in retroactive effects of memory. *Journal of*

Experimental Psychology: Learning, Memory, and Cognition, 44(9), 1352–1364. https://doi.org/10.1037/xlm0000512

Neisser, U. (1982). Literacy and memory. In: U. Neisser (Ed.), *Memory observed: remembering in natural contexts,* Chapter 23, 241-242. New York: W.H. Freeman.

Nettl, B. (1983). *The study of ethnomusicology: Twenty-nine issues and concepts* (No. 39). University of Illinois Press.

Palmer, C., & Pfordresher, P. Q. (2003). Incremental planning in sequence production. *Psychological Review, 110,* 683-712.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, *77*(3p1), 353.

Posner, M. I., & Keele, S. W. (1970). Retention of abstract ideas. *Journal of Experimental Psychology*, *83*(2p1), 304.

Postman, L. (1972). A pragmatic view of organization theory. In E. Tulving & W. Donaldson, Organization of memory. Oxford, England: Academic Press.

Quesnel, R. (1996, May). Timbral ear trainer: Adaptive, interactive training of listening skills for evaluation of timbre. In *Audio Engineering Society Convention 100*. Audio Engineering Society.

Quesnel, R. (2002). *A computer-assisted method for training and researching timbre memory and evaluation skills* (pp. 1-131). Doctoral thesis, McGill University.

Rosch, E. (1999). Principles of categorization. In E. Margolis & S. Laurence (Eds.), *Concepts: core readings*, pp. 189-206. MIT Press.

Rubin, D. C. (1995). *Memory in oral traditions: The cognitive psychology of epic, ballads, and counting-out rhymes*. New York: Oxford University Press.

Rumelhart, D. E. (1975). Notes on a schema for stories. In *Representation and understanding* (pp. 211-236). Morgan Kaufmann.

Salamé, P., & Baddeley, A. (1989). Effects of background music on phonological short-term memory. *The Quarterly Journal of Experimental Psychology*, 41A, 107-122.

Savage, P. E., Brown, S., Sakai, E., & Currie, T. E. (2015). Statistical universals reveal the structures and functions of human music. *Proceedings of the National Academy of Sciences*, *112*(29), 8987-8992.

Schank, R.C. & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding: An inquity into human knowledge structures.* New York: Psychology Press.

Schellenberg, E. G., Iverson, P., & Mckinnon, M. C. (1999). Name that tune: Identifying popular recordings from brief excerpts. *Psychonomic bulletin & review*, *6*(4), 641-646.

Schulze, K., Dowling, W. J., & Tillmann, B. (2012). Working memory for tonal and atonal sequences during a forward and a backward recognition task. *Music Perception, 29*(3), 255-267.

Schulze, K. & Koelsch, S. (2012). Working memory for speech and music. *Annals of the New York Academy of Sciences, 1252*(1), 229-236.

Segalowitz, N., Cohen, P., Chan, A., & Prieur, T. (2001). Musical recall memory: Contributions of elaboration and depth of processing. *Psychology of Music*, *29*(2), 139-148.

Shepard, R. N. (1995). Mental universals: Toward a twenty-first century science of mind. *The Science of the Mind: 2001 and Beyond*, 50-62.

Shepard, R. N. (2009). One cognitive psychologist's quest for the structural grounds of music cognition. *Psychomusicology: Music, Mind and Brain*, *20*(1-2), 130.

Shepard, R., & Levitin, D. (2002). Cognitive psychology and music. In D. J. Levitin (Ed.), *Foundations of cognitive psychology: Core readings*, 503-514. Cambridge, MA: MIT Press.

Sloboda, J. A., Hermelin, B., & O'Connor, N. (1985). An exceptional musical memory. *Music Perception: An Interdisciplinary Journal*, *3*(2), 155-169.

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, US: Lawrence Erlbaum Associates, Inc.

Tirovolas, A. K., & Levitin, D. J. (2011). Music perception and cognition research from 1983 to 2010: A categorical and bibliometric analysis of empirical articles in Music Perception. *Music Perception: An Interdisciplinary Journal, 29*(1), 23-36.

Tooby, J., & Cosmides, L. (2007). Evolutionary psychology, ecological rationality, and the unification of the behavioral sciences. *Behavioral and Brain Sciences*, *30*(1), 42-43.

Tse, A. & Levitin, D. J. (2019). Memory for absolute loudness. *Proceedings of the Annual Meeting of the Society for Music Perception and Cognition, New York, NY.*

Tymoczko, D. (2006). The geometry of musical chords. *Science*, *313*(5783), 72-74.

Tyng, C. M., Amin, H. U., Saad, M. N., & Malik, A. S. (2017). The influences of emotion on learning and memory. *Frontiers in psychology*, *8*, 1454.

von Ehrenfels, C. (1890). Über Gestaltqualitäten [On Gestalt qualities]. *Vierteljahresschrift für wissenschaftliche Philosophie, 14*, 249-292.

Voss, R. F., & Clarke, J. (1975). 1/f noise in music and speech. *Nature, 258*, 317–318.

Wachter, J. A., & Kahana, M. J. (2019). *A retrieved-context theory of financial decisions* (No. w26200). National Bureau of Economic Research.

Walden, C., Berg, S., Reyes, C. (2019). Personal communication, July 12, 2019. Sonoma County, California.

Ward, W. D., & Burns, E. M. (1978). Singing without auditory feedback. *Journal of Research in Singing*, *1*(2), 4-44.

West, R., Howell, P., & Cross, I. (1985). Musical structure and cognition.

White, B. W. (1960). Recognition of distorted melodies. *The American Journal of Psychology, 73,* 100-107.

Wu, D., Kendrick, K. M., Levitin, D. J., Li, C., & Yao, D. (2015). Bach is the father of harmony: revealed by a 1/f fluctuation analysis across musical genres. *PloS one*, *10*(11), e0142431.