Musical Affordance and the MP3
Boundary 2 43/4 (2016)

Richard Randall is the director of the Music Cognition Lab at Carnegie Mellon University where he is faculty of the School of Music and the Center for the Neural Basis of Cognition. He is the co-director of Listening Spaces—a three-year project investigating the role technology plays in everyday music making.

Abstract. The materialization of music by means of notation and recording has had a number of consequences. On one hand, it facilitates the transmission of music across space and time. On the other hand, formats have allowed music to become an excludable and rivalrous good. Formats embody a tension between music as an emergent social practice and music as means for controlling social practice. Jonathan Sterne’s recent book, MP3: The Meaning of a Format, explores the subject of digital audio formats in general and the MP3 in particular. This essay problematizes Sterne’s ideas of perceptual coding, compression, and format in an attempt to place them in a broader cultural and psychological context. By doing so, I hope to provide a clearer picture of who we are as musical beings and offer an alternate view of how musical formats serve us. I refine Sterne’s format theory to include the idea of affordance and argue that the MP3 facilitates a new era of participatory media.

Christopher Small coins the term “musicking” as a verb that describes a broad collection of activities that comprise musical engagement. He proposes that being musical involves not just performing and creating, but also listening and sharing. Most of us find ourselves doing the latter two more often than the former, and such activities necessarily require that the acoustic manifestation of music be recorded and stored in a particular format to be realized again later. Such formats, it turns out, have become an integral part of our musical lives and relate to us in important ways. The materialization of music by means of notation and recording has had a number of consequences. On the one hand, it facilitates the transmission of music across space and time. I can share music from my hometown with people thousands of miles away and I can listen to performances of long dead musicians. On the other hand, formats have allowed music to become an excludable and rivalrous good. Excludability means that one person can prevent another person from having access to a good unless they have paid for it. Rivalry ensures that if one person owns a good, then another person cannot. Formats embody a tension between music as an emergent social practice and music as means for controlling social practice. Jacques Attali writes that “music, an immaterial pleasure turned commodity, now heralds a society of the sign, of the immaterial up for sale, of social relation unified in money.” There is little question that the materialization of music has had a profound influence and effect over musical practice, especially in the West. Attali argues that we devised musical formats to allow us to exercise political control over what music is and how it can be used. “Wherever there is music,” he says, “there is
money" (Noise, 3). The spectacle of music lies in the consolidation of political power and as a representation of social and economic control. Formats play a critical role in this consolidation. Attali writes:

“Music is more than an object of study: it is a way of perceiving the world . . . Fetishized as a commodity, music is illustrative of the evolution of our entire society; deritualize a social form, repress an activity of the body, specialize its practice, sell it as a spectacle, generalize its consumption, then see to it that it is stockpiled until it loses its meaning. Today, music heralds—regardless of what the property mode of capital will be—the establishment of a society of repetition in which nothing will happen anymore” (Noise, 5).

However, Attali imagines a way out: “. . . at the same time, it heralds the emergence of a formidable subversion, one leading to a radically new organization never yet theorized, of which self-management is but a distant echo” (Noise, 5). He envisions “not a new music, but a new way of making music” that exists outside of the signs and codes of society (Noise, 134). We often understand our musical practice in terms of oppositions: music/noise; musician/non-musician; performer/audience; producer/consumer; owner/user. Many of us do not know music as anything else. Musicking is a powerful idea because it dislodges musical practice from specific power systems (e.g. the conservatory, the concert stage, the court, the record company) and restores the idea of music as a social form and activity of the body. It gives everyone permission to call themselves a musician (or, perhaps more accurately, a musicker) and allows us to reclaim specific actions as being musical, circumventing sanctioned systems of specialization, production, and ownership. Actions such as sampling, mashups, remixing, and file sharing subvert the systems of power that Attali describes because to Small they are all legitimate musical practices. Of course, these actions exist within the overarching system and therefore use the tools and materials provided for by the system. The musician becomes a hacker of sorts, using tools and materials in ways the system never intended. And the tool that has the greatest recent impact on how we experience music is the digital audio format.

Jonathan Sterne’s recent book, MP3: The Meaning of a Format, explores the subject of digital audio formats in general and the MP3 in particular. Sterne is not the first author to discuss MP3 or issues of digital audio formats in a cultural context, but the scope and depth of his book distinguishes it as a unique and significant contribution to the sound-studies corpus. His documentation of the history and development of the MP3 positions audio formats as a worthy subject of study and something that we often take for granted. Sterne writes that “more (music) recordings exist and circulate in MP3 format than in all other formats combined” (MP3, 1). The MP3 phenomena is simply too large and has been too transformative to ignore. Humans have been explicitly making music for over 40,000 years. We developed the technology to record and play back music about 135 years ago.
The computer code for the MP3 was developed about 25 years ago. Today, Sterne estimates that 1 to 1.5 billion music files circulate every month between 10 million users on just one of a vast number of peer-to-peer file-sharing networks (MP3, 1).

The history of musical formats, however, is neither as positive nor teleological as Sterne seems to suggest. We are not led from practice to practice in ever increasing meaning, importance, and relevance for society. Sterne’s account of the creation of the MP3 provides a tremendous amount of detail, but is largely uncritical in terms of how formats impact music as a cultural practice. Critics such as Attali and Peter Szendy demonstrate that musicking today is done within a capitalistic system that relies on the enforcement of differences in order to maintain control and profitability. The MP3 is unique in that it not only has the capacity to maintain established values of simplification and consolidation for the sake of efficient control and commercialization, but also has the capacity to undermine them. If we take as axiomatic our need for community and that music facilitates this, then the question is: What roles do Sterne’s formats and compression play? Before any kind of recording and playback technology, participating in music meant making music or seeing music performed. Today, participation takes on many forms and the success of the MP3 lies in its ability to let us do things with music that no other format could do.

Sterne presents a focused narrative that leads from the invention of sound recording and reproduction to the present day. He locates the MP3 in values and techniques established by the telecommunications industry at the turn of the century. Those values revolved around maximizing the intelligibility of human speech as it is converted into an electric signal, transmitted via wire cable, and finally converted back into an audio signal at the other end of the line. Sterne is convincing when he says that the MP3 embodies our technological and economic acumen. More controversial is his claim that the MP3 also embodies our physiological and psychological makeup. The impetus for this position is the perceptual coding used in the compression algorithm that makes MP3s. What is missing from Sterne’s account, however, is a broader understanding of who we are as musical beings and the role this identity plays on the development and use of digital audio formats (including the MP3).

Sterne articulates a need for a format theory that would explore “a whole range of decisions that affect the look, feel, experience, and workings of a medium” (MP3, 7). There is little doubt that this is a much-needed addition to critical media and sound studies. He argues that formats matter in terms of their “contexts of their reception, conjunctures that shaped their sensual characteristics, and the institutional policies in which they were enmeshed” (MP3, 11). He views the MP3’s story as an excellent starting point for a variety of reasons. First, its transformative presence is palpable. Many of us can remember the sudden ability to quickly acquire almost any song we could imagine for “free,” and the legal drama that unfolded in its wake. Many of us now struggle to remember a life before the Internet and the ubiquity of being connected. Second, the MP3’s development is surprisingly well
documented. It is a format created by a group whose research and testing was purposefully organized and recorded (MP3, 134).

MP3 is an encoding standard used to make space-saving versions of digital-audio files. More specifically, MP3 is a codec (a portmanteau of coder-decoder), which is a set of instructions for writing and reading a data file. “MP” is short for Moving Picture Experts Group (“MPEG”) and refers to the governing body that established digital encoding standards in the 1980s. Over the years, this group has defined a number of digital video and audio standards for the purposes of helping industry “offer end users an ever more enjoyable digital media experience.”

The “3” indicates the particular version of the MPEG code (called a “layer”) that uses a psychoacoustic model to remove details it believes a listener will not miss. The MP3 is a member of a family of compression standards for storing and transferring audio and video, all of whom represent patented proprietary technology. The MP3 came to dominate this family through an unparalleled confluence of events, actions, and technology – in part what Sterne calls “transectional innovation” (MP3, 203). In the mid-1990s, the proprietary code used to make an MP3 was hacked, repackaged, and freely distributed (MP3, 201). Simultaneously, cultures of music sharing and piracy blurred as access to broadband technology and personal computers outpaced the recording industry’s ability to reify ambiguous notions of materiality and ownership of digital goods. Soon, millions of users latched on to MP3 as a way to encode their music collections for online sharing and trading. Rather than incorporating this technology into a new business model, the record industry’s main strategy was to litigate and criminalize online sharing. They soon lost their footing as peer-to-peer file sharing spun out of their control. The introduction of portable media players such as the MPMan (1997) and the Diamond Rio (1998) and CD-Rs spurred the revolution by making MP3 playback part of a mobile lifestyle. Over the next decade, storage capacity of these devices increased to the point where individuals could carry their entire music collection on a device the size of a deck of cards.

Sterne describes the MP3 as a format capable of telling a specific kind of story of who we are as musical beings. The algorithm that allows an MP3 to represent a larger music file as a smaller, yet extremely similar version, is based on psychoacoustics – the study of how we perceive sound. Psychoacoustics, a subfield of psychophysics, is concerned with low-level perceptual discriminations such as the range of human hearing, the minimum duration a sound event must have in order for us to perceive it as a pitch, and how one sound can mask another when they are played simultaneously. The MP3 takes these kinds of details into account. Perceptual coding means that perceptually unimportant information is cast aside, while important information is retained (MP3, 92-127). The overriding idea being: if you can’t hear it, then you won’t miss it. The ideal application of the coding algorithm creates a much smaller version of the original that is perceptually indistinguishable from the original – a crucial issue when digital storage and bandwidth are an issue. This equivalence is not as important in practice as it is in theory, however. Sterne does a good job of describing these issues in language that a general audi-
ence can understand. More importantly, Sterne takes time to provide a critique of psychoacoustics as a means to understand human hearing (MP3, 241). Psychophysical studies, in general, allow us to make extremely precise observations of extremely limited scope. They often involve running a small number of subjects through thousands of trials in an attempt to identify precise thresholds of perception (the just-noticeable difference), and results often cannot be generalized across a population. In addition, such studies suffer from a lack of ecological validity—they test how we hear man-made sounds in contrived man-made environments. Scientists working on compression algorithms understood this and developed a second round of experiments that were used to fine-tune the MP3 codec. These experiments worked to answer the question: Is the MP3 encoding a good approximation of the original uncompressed digital file? They employed studies where people listened to actual musical recordings under different versions of the MP3 algorithm in an attempt to find the one that best represented the original. This raises interesting questions about the guise of scientific objectivity once we move “higher up” beyond the level of psychoacoustics. In an interview with one of the scientists responsible for such tests, Sterne finds that the choice of music used was due to the tastes of the experimenter (MP3, 169-171). Sterne writes:

“(H)e speaks of music not for its universality but for its effect on him. His affection for some of this music—and music in general—came through clearly in his answer. This is one of the central contradictions of the listening test. It operates according to a logic and rhetoric of disinterest, yet it is equally susceptible to what Pierre Bourdieu called “the games of taste” . . . The listening test presents the technology as if it could work for anyone; but by virtue of playing music, its sonic address is always to and for someone in particular” (MP3, 171).

To listen to an MP3, in other words is to listen to music through the ears and tastes of others. This observation is especially poignant in light of my previous discussion of the tension between music as social practice versus as a material commodity. This is reminiscent of Peter Szendy’s discussion of the idea of the musical arrangement, which I discuss in more detail later. Szendy writes that, “(b)eyond its social or public function of communication and diffusion (the work for orchestra is supposed to circulate more easily in a piano reduction), we should analyze other functions of arrangement, such as those we might call clarifying or corrective.” Szendy’s terms “clarifying” and “corrective” can be read in a number of ways, as we will see, but the important connection is that Szendy’s arranger functions similarly to Sterne’s scientist in that they both exert influence on the final product that the consumer hears. In other words, this practice is not new. In the digital era, this influence has become a suppressed premise and Sterne commendably raises it to a supraliminal level.

We can usually find an MP3 version of an uncompressed digital audio file that is indistinguishable from the original. But interestingly, this equivalence is not as
important to listeners in practice. In mid-2012, computer scientist and blogger Jeff Atwood created an online survey on bitrate preference for the MP3 format. Bitrate is the metric that describes how many thousands of bits of information (kilobits) are represented within one second (kbps). MP3 can encode music in a wide range of bitrate resolutions. The higher the kbps, the more likely you are to accept the copy as the original; the lower the kbps, the more likely you are to hear distortions and artifacts. Of course, the higher the resolution, the larger the file is. Encoding an MP3 is about balance and compromise. The larger the file, the fewer you can store, and the longer it takes to upload or download.

In Atwood’s investigation, participants were challenged to listen to five versions of Starship’s “We Built the City (on Rock ‘n Roll)” and rate each clip on a scale of one to five. This informal experiment does not meet scientific standards, but it is nevertheless informative. The results showed that people could easily single out the example encoded at a low rate of 128kbps. What was surprising is that by increasing the bit rate slightly listeners could not distinguish music encoded at 160kbps from music encoded at the maximum 320kbps. Of course, no matter what bitrate we are listening to, we will always be constrained by the digital-to-analog converter, the speakers, and the ambient environment. All of which goes to show that the quality of sound we require is dependent on where, when, how, and why we are listening. Sterne acknowledges that it is not that the MP3 got it right, but rather that it got it good enough. The website Soundexpert.org is also in the testing business. They have been continually running experiments on listener preference not just for different bitrates, but also for different encoders and portable playback devices. They concluded that although people can distinguish between different encoders or portable players, “…most people don’t care much about sound quality they hear from stereos, computers, and MP3 players. To be honest, Music itself is so great it can impress even through poorly sounding equipment.” This supports the idea that when it comes to ideas of resolution, most people are willing to compromise if the format affords them the ability to carry more music with them or listen to a song through a variety of personal computers and digital devices.

Perceptual coding is not new, nor is it limited to the domain of digital audio. In fact, the way we listen to music allows us to focus on what psychologists call “invariants in perception.” These are the details that we consider the same even when other factors have been changed. We can easily identify a familiar tune (e.g., “Happy Birthday To You”) regardless of the instrumentation or key. In many ways, “Happy Birthday” is the same to us played on a violin or a kazoo. Of course, we can point to differences between two such realizations, but we can also accept the two as being the same. When we consider music as an intrinsic part of a social context (e.g., the birthday celebration), how it is realized becomes much less important than the fact that it is realized. This is a different kind of perceptual coding that Sterne does not discuss, but is nevertheless critical to understanding the degree to which music is a part of our lives and the role formats play. The tune “Happy Birthday” is an interesting example because it represents a social form that has been
created by a system of power: the publishing and copyright legal infrastructure. Easily one of the most recognizable songs in the English language, it is owned by a subsidiary of AOL Time Warner. There is no question that the ubiquity of the song is due to its publishing and wide distribution beginning in the early twentieth century. However, it is part of our collective consciousness, and has been fully incorporated into recurring social rituals. The current copyright holder is far removed from original authors of record, who may or may not be the source of the song. Robert Brauneis writes:

“Moreover, many have suggested that, notwithstanding the attribution of the song to the Hill sisters, it is so much like other previous songs that it should be treated as having arisen from a folk tradition rather than the creative talents of a particular author. “Happy Birthday to You” is not only currently under copyright, but will supposedly be under copyright until the year 2030 – 137 years after 1893, an incredibly long time even by the standards of the Copyright Term Extension Act . . .”

We use the song at will within the ritual of celebration, and the notion of ownership in this context paints a grotesque caricature of the very power system at the heart of Attali’s critique.

The methods of psychophysics and behavioral psychology would seem to fall prey to the criticisms Attali levies against the use of materialized music. Psychology often posits simplified models of human behavior in an attempt to describe who we are, what we do, and why we do it. The practices that lead to these seemingly reductionist statements are highly restrictive and are designed to consolidate power into structures and systems that are authorized to carry out research. These practices have been critiqued most notably by Paul Feyerabend. He takes the position that if scientific inquiry were to truly conform to the scientific method, then we would be forever cut off from the creation of new knowledge. Nevertheless, the study of music from social, structural, psychological, and anthropological perspectives can elucidate what music is, why we are musical, and how musical we really are. These perspectives can help us discuss Attali’s idea of music as a social form and activity of the body more specifically. It is beneficial, in the case of Sterne’s thesis, to look more deeply at both what scientific and cultural studies can tell about how and why we are musical. We are, in fact, very musical as a species and an understanding of our musical abilities can help contextualize compression in a broader sense. Infants as young as three-months old prefer consonant chords and music to dissonant chords and music. Studies show that eight-month-old infants are equally good at telling when a melody contains a wrong note, whether it is out of the key, or out of the implied harmony. Five-year-olds hear according to key structures. At about seven years of age, children become sensitive to implied harmony. Adults with no formal musical training have learned a lot about musical structure through years of exposure through a principle called “statistical learning.” Given a musical sequence of tones or chords, we can identify “wrong” notes in a
melody when they go outside the key or implied harmony of the melody. Non-musicians can remember words to songs and melodies just as easily as trained musicians. We can identify whether or not two melodies are the same, even if they are played at different tempos, on different instruments, and in different modes and keys. We all possess these abilities. To eschew the common musician/non-musician dichotomy for a moment, we are all musicians to a very high degree. This kind of learning is culturally based and exists in a feedback loop with Attali’s power systems. We are musical, but how we are musical depends on what is available to us. The extreme consolidation of musical styles in any given culture, for example, increases the efficacy of statistical learning.

Using a music-cognitive perspective, we can broaden the ideas of compression and format that Sterne describes. As a simple example, let us consider Beethoven’s 5th Symphony, which was premiered in 1808. While not immediately heralded as a masterpiece, it gained such status after the publication of the score. For those who read E. T. A. Hoffmann’s glowing review, the only opportunity to experience this magnificent piece would have been to see a live performance or to acquire the full score and imagine the sounds in your mind’s ear. The format in this example is the score, which shows the notes of each instrument in the orchestra individually. Realizing the symphony from the score in your mind’s ear would require a level of training and expertise that few people have. However, many people that might have been interested in this piece would likely have owned and been able to play a piano. In 1838, Franz Liszt published his now famous piano (in four hands) transcription of the 5th Symphony, making the experience of the piece widely available. This transcription is a compressed version of the original. Its realization requires only two people and one instrument (a piano) as opposed to the original, which required 100 musicians playing 17 different kinds of instruments. The piano score was a format that was easily readable by amateur musicians. In this example, the musical qualities that are preserved are the structural interrelationships of the sounds themselves. The complexity of the original is reduced from 100 people playing 17 kinds of instruments to a version with two people playing a single instrument. Peter Szendy describes this act as “arranging,” although we can just as easily call it “compression.” We accept that the arrangement is a valid representation of the original even though we would never mistake one for the other. Some of the musical information is lost, but we are able to accept this substitution because the reduction affords us things that the original does not. The reduction lets us bring music into our home, it affords us the experience of a new and widely praised piece of music, it affords us the excitement of being a performer, and affords us the opportunity to bring people together as performers and listeners.

With a basic understanding of who we are as musical beings, we can focus on social aspects of musicking. Daniel Pinker asserted that music was simply “auditory cheesecake.” For Pinker, this meant that music is a parasitic byproduct of other physical and mental capacities developed in direct response to specific pressures. The analogy is that just as cheesecake sates our drive to consume calories necessary for survival, we could have survived equally well without it. We don’t
need cheesecake; we have it because we like it. Likewise, for Pinker, music is a spinoff of selective pressures such as language development, auditory scene analysis, and emotional calls, but is not itself necessary for survival. Steven Mithen, however, provides a convincing account that not only was it necessary for our (Homo sapiens) survival, but it was also necessary for the survival of Homo neanderthalensis and many of our evolutionary ancestors. His argument is that before language, there was a protolanguage that was holistic, multi-modal, manipulative, mimetic, and musical. As pressures increased, language became representational and propositional; music became emotional and focused on building bonds. Music remained because it did these things effectively and efficiently. Erik Clarke writes that “music affords dancing, worship, coordinated working, persuasion, emotional catharsis, marching, foot-tapping, and a myriad other activities of a perfectly tangible kind.”

Sterne acknowledges this when he writes: “in recent years, social psychologists and ethnomusicologists have moved from the expectation that their subjects treat music as an end in itself to the presupposition that listening is usually part of a broader sphere of social activity” (MP3, 5). This has its roots in distinction between “participatory music” as opposed to “performance music” (or “audience-oriented music”). The mistake in Western-music scholarship from the 19th century onward was to confuse the latter with the former. Small’s idea of musicking challenges this distinction. As we shifted away from participatory music and toward the more passive listening modes, we went through great lengths to get music into our homes. Before sound recording and reproduction, there were court musicians, songbirds, piano transcriptions, player and reproducing pianos, and automatic mechanical musical instruments. All of these technologies allowed people to participate in music. The MP3 facilitates a new era of participatory media.

Joel Krueger’s 2011 article “Doing Things with Music” provides an excellent treatment of musical affordance. Challenging ideas that listening and musical experiences are passive, Kruger argues that music is something we are always seeking. Music, Krueger writes, “is a crucial tool for cultivating and regulating our social life. Without music, our life—including our ability to sensitively relate to and communicate with others—would indeed change dramatically.” As mentioned above, if I can fault Sterne for anything, it is for interpreting perceptual coding as a hidden code in the format that tells the story of “practical and philosophical understandings of . . . what it means to make music” (MP3, 2). Formats do tell stories about what it means to make music, but this refers to what kind of musical activities they afford us to do. The history of musical formats is filled with ways humans have represented music in order to do something specific with it. Each format allows us to do something different, and formats live or die by the activities they afford. Here we have two important questions: First, what is it that music affords us? Second, how do formats support these affordances? To begin answering this question we must understand that music is about us making music both individually and collectively. Erik Clark writes that “in certain musical traditions (and the concert music of the
West is an obvious example) listening to music has become somewhat divorced from overt action. Sterne rightly reveals a general misconception that audio formats follow a monotonically positive trajectory of verisimilitude (MP3, 4). We have a sense that the “more truthfully” a format can represent a stimulus, the better it is, and that the goal of a format is to move us toward this auditory perfection. Sterne raises doubts about the idea that “increases in definition necessarily enhance end-user experience” (MP3, 4). His conclusion invites the question as to why this is. A live rock-music performance is as real as it gets, but hardly sets a standard for fidelity. Rock clubs are noisy places with people talking, bad acoustics, and poor sound reinforcement. However, we do not go to these events for sound fidelity. We go to see people create music spontaneously. We go to verify that the people making the music that we care about are real. We go to be with people who share our values. We go to exhibit behaviors that reflect the values of a particular social group. We go to participate and show solidarity. We go to accept and to be accepted. This is what Judith Becker calls “habitus of listening.” Becker writes, “A habitus of listening suggests . . . an inclination, a disposition to listen with a particular kind of focus, to expect to experience particular kinds of emotions, to move with certain stylized gestures, and to interpret the meaning of the sounds and one’s own emotional response to the musical event in somewhat (never totally) predictable ways.”

One thing that music has done for us as a species is brought us together. Together, we have been able to survive. Making a social group stronger by making music together is one of the most important roles that music has played in our history. Attali positions this one of the many ways that music supports politicization, and I offer that the benefice of this practice outweighs the negative (and perhaps inevitable) developments of materialization, ownership, and suppression seen throughout the 15th century and on in the West. We care about music, in part, because we care about others and ourselves. We recognize that the togetherness afforded by music increases our chances of survival. Today music is just as important to us as it has always been. Many details have changed in the last 20 years; the most important of which are the ways we find, share, listen to, and store music. Today’s technological advances in digital communication and storage have facilitated new communities built around music. Invitation-only torrent-tracker sites are just one example of groups of people coming together to build online communal libraries for sharing. These communities have rules for membership to ensure the quality of the shared music and social civility. These communities exist because music is available in a format that resonates sympathetically with who we are and how we live our lives. We no longer need to live in the same country with another person in order to develop a relationship with them. Subject-specific online forums allow people to congregate with each other over shared interests while never requiring them to physically be in the space place. As I mentioned earlier, the success of the MP3 lies in its ability to let us do things with music that no other format could do. According to Attali, star performers and producers are as important to the political economy of music as merchants and publishers. Without them, there would
be no product to sell and they work together to systematically regulate music. The MP3 can be credited with creating a new star system for music. Shawn Fanning (founder of Napster), Kim Dotcom (founder of Megaupload), and Alan Eliss (founder of OiNK), as just three examples, are all famous “musicians” for anti-authoritarian reasons. By facilitating musical activities of listening and sharing, they are musicking. Their star status was earned by providing the means by which Internet users could share music with others completely outside of established and approved channels. Each of them also fell to a system of power that viewed their contributions as enabling illegal distribution of copyright-protected music. Musicking is an idea that is morally and ethically agnostic and practices such as “piracy” and “stealing” can be construed to be as musical as “composing” and “concertizing.”

This essay problematizes Sterne’s ideas of perceptual coding, compression, and format in an attempt to place them in a broader cultural and psychological context. By doing so, I hope to provide a clearer picture of who we are as musical beings and offer an alternate view of how musical formats serve us. I refine Sterne’s format theory to include the idea of affordance. This idea appears occasionally throughout Sterne’s book in unspecific ways, and makes a brief appearance in his discussion of the ontology of music. “If music is a thing,” Stern writes, “then it is for something” (MP3, 193). Since musical formats present alternate representations of the same thing (i.e. “music”), their differences lie in what each format (i.e. re-representation) offers us in contrast to the original or alternatives or in how it allows us to interact with each other.

Sterne says, “The history of the MP3 belongs to a general history of compression” (MP3, 5). This, I believe, is his most profound idea. The history of compression has yet to be written but Sterne gives us a good idea of what it might look like. For Sterne, audio compression is synonymous with lossy compression and perceptual coding. Sterne says that “whatever happens to MP3 as a format, compression has a long future ahead” (MP3, 230). There are, however, other ways to think about compression in the context of music, and we can use compression to make general observations about who we are psychologically and how we use music in our everyday lives. From this point of view, we see the MP3 as means of allowing to people to participate in music in a radically new technological milieu with powerful social implications. More music is being shared today than at any other point in history. Today, more people have more access to more music then ever before and the MP3 remains the seminal format for online sharing, trading, blogging, developing communities, and facilitating connections between people.

Digital music formats have created a tension between traditional modes of music production and consumption. The consolidation of music into highly controlled distribution and production networks has allowed specific systems to exploit and profit from what can be considered an irrepressible human phenomena. The ubiquity of the MP3 is due the role music plays in our lives. Musicality needs to be understood as an inalienable right, not a legal right. Formats were created as tools of power to regulate and profit from an ongoing social practice. In the West, it irrevocably changed our concept of music. The transectional innovation of the MP3
and the Internet creates an environment where the very notion of a format as a tool of economic and political control is called into question. To rephrase Margret Grebowicz’s discussion of pornography in the Internet age: the question as to whether musical formats are hegemonic (part of the problem, as Attali would suggest) or transgressive (part of the solution, as I am suggesting in the case of the MP3) depends on how easily an individual or community is able to act with agency in order to perform simple inalienable musical actions. The MP3 has been an unqualified success simply for the fact that it has reminded us that we are all musical.

7 Coleman, *Playback: From the Victrola to MP3, 100 Years of Music, Machines, and Money*, 10.
8 Sterne associates these figures to the Gnutella network at the time of his writing. Given the legal questions surrounding sharing digital audio files online, the numbers of active members or files being shared can fluctuate wildly should one or more of these networks be shutdown. Take for example the file-sharing site Megaupload.com, which recorded over 50 million visitors per day until it was shutdown in January 2012.
9 As with much media, cost is often totally or partially deferred to the equipment used to acquire and playback that media.
MPEG has three layers, each with different properties. Layers 2 and 3 both use a psychoacoustic model for compression. Of these two, MP3 is the most popular for Internet applications, whereas MP2 remains popular in commercial usage. See J. S. Chitode, Information Coding Techniques (Pune: Technical Publications, 2007).

This is one example of a lossy compression algorithm. Lossy compression formats are those that discard information that the algorithm deems unnecessary. Once discarded this information can never be retrieved and is considered “lost.” Lossy formats are contrasted against lossless formats, which while making a file smaller, can always be uncompressed back into their original format with no information loss.


Clarke, Ways of Listening, 32.


Brauneis, “Copyright and the World’s Most Popular Song,” 2.

Paul Feyerabend, Against Method (New York: Verso, 1993)

Isabelle Peretz, “The Nature of Music from a Biological Perspective,” Cognition 100: 1–32. These are three of the four fundamental questions of music cognition identified by Peretz.


27 Franz Liszt, Beethoven Symphonies 5 and 6, arr. Piano 4-Hands (Leipzig: Breitkopf and Härtel, 1838).
28 The standard for a full-size orchestra.
30 Stephen Mithen, The Singing Neanderthal: The Origins of Music, Language, Mind, and Body. The Neanderthal line ultimately died out around 30,000 years ago after surviving for over 150,000 years. There is scant evidence that the Neanderthals had the capacity for symbolic though and language as we know it. There is, however, considerable evidence that this species used a sophisticated protolanguage that resembling Brown’s “musiclanguage.” The musical function of this language (social grooming, emotional expression, sexual selection, bond building) is hypothesized to have played a crucial role in the survival of both of our species.
31 Clarke, Ways of Listening, 38.
35 Clark, Ways of Listening, 38.
37 Torrent tracker sites such are What.cd and RevolutionTT are exclusive communities for trading a variety of media formats.