

## AUDITORY SCENE ANALYSIS AND THE PERCEPTION OF SOUND MASS IN LIGETI'S *CONTINUUM*

---

CHELSEA DOUGLAS, JASON NOBLE &  
STEPHEN MCADAMS  
*McGill University, Montréal, Québec, Canada*

SOUND MASS HAS BEEN AN INFLUENTIAL TREND IN music since the 1950's and yet many questions about its perception remain unanswered. Approaching sound mass from the perspective of auditory scene analysis, we define it as a type of auditory grouping that retains an impression of multiplicity even as it is perceived as a perceptual unit. Sound mass requires all markers of the individual identities of sounds to be deemphasized to prevent them from splitting off into separate streams. Seeking to determine how consistent listeners are in their perception of sound mass, and whether it is possible to determine sound parameters and threshold values that predict sound mass perception, we conducted two perceptual studies on Ligeti's *Continuum*. This piece consists of an extremely rapid, steady stream of eighth-note dyads with no tempo changes. We addressed the claim by Ligeti and others that the fusion into a continuous texture or sound mass occurs at ca. 20 attacks/s, hypothesizing that other factors such as pitch organization, emergent rhythm, timbre, and register would affect this value. A variety of factors were found to affect sound mass perception, suggesting that the threshold value is not absolute but varies according to principles of auditory scene analysis.

*Received: December 7, 2014, accepted April 15, 2015.*

**Key words:** sound mass, fusion, auditory scene analysis, Ligeti, *Continuum*

---

**S**OUND MASS EXISTS WHEN THE INDIVIDUAL identities of multiple sound events or components are attenuated and subsumed into a perceptual whole, which nevertheless retains an impression of multiplicity. Typically this involves one or more parameters of sound — for example, rhythmic activity, pitch organization, or spectral content — attaining a degree of density, complexity, and/or homogeneity that is perceived as saturation. This phenomenon may be observed in the

perception of many everyday sounds such as shattering glass, swarming insects, and rustling leaves. The term appears frequently in the musicological literature to describe a musical trend originated in the 1950's–1960's by Witold Lutosławski (1913–1994), Iannis Xenakis (1922–2001), György Ligeti (1923–2006), and Krzysztof Penderecki (b. 1933), and later explored by many 20th and 21st century composers.

Although this approach to composition has been highly influential in the decades since its inception, many questions about sound mass remain unanswered from a perceptual point of view. To what extent do listeners agree on whether or not a particular stimulus is a sound mass? Is it possible to identify specific musical parameters and values that define the threshold of sound mass perception? Is the type of synthesis at work in sound mass perception different from other types of perceptual fusion such as harmonic integration and timbral blend?

There is currently no universally accepted definition of sound mass. Several authors (e.g., Bernard, 1999; Edwards, 2001) have defined it as a de-emphasis of pitch as a privileged sound quality, such that other parameters such as texture, timbre, and dynamics become the foci of musical attention. However, we believe that sound mass perception requires not only pitch but *all* markers of the individual identities of sound events to be deemphasized, including their dynamic contours, rhythmic profiles, and timbre saliency. Otherwise, those markers could cause some sounds to segregate into separate auditory streams, disintegrating the perception of a singular mass.

The many factors related to auditory stream integration and segregation are detailed in Bregman's *Auditory Scene Analysis* (1990). In that seminal text, Bregman describes how a variety of perceptual principles, such as fundamental frequency (the perception of identifiable pitch in a given sound, related to the acoustical phenomenon of harmonicity) and common fate (correlated trajectories of frequency or amplitude of simultaneous sounds or sound components) allow the auditory system to parse simultaneous incoming vibrations into separate real or virtual sound sources. Other principles such as temporal proximity (the durations of gaps

between subsequent sound events), loudness, and spectral shape (relative amplitudes of partials over time) determine the integration of temporally extended auditory streams.

Sound mass, as we have defined it, is closely related to auditory streaming in that it involves the perceptual grouping of incoming auditory stimuli. However, an auditory stream is typically perceived as emanating from a single sound source, whereas sound mass retains an impression of multiplicity: it is a totality rather than a singularity, many coming together into one. And yet it is not simply the auditory whole, an unqualified assimilation into which all sounds and sound components in the auditory field are subsumed. It may be possible to perceive multiple sound masses simultaneously, for instance, a roaring fire and a flock of seagulls, or a crowd of people and traffic noise. Also, sound mass may coexist in the perceptual field with other, segregated streams, as when we perceive the ambient noise of a crowd as a sound mass but nevertheless distinguish the voice of the person next to us in conversation. Sound mass fusion is similar to the orchestration technique of textural integration (McAdams, 2014) in which multiple instruments playing independent musical lines are nonetheless integrated into a global texture. In sound mass perception, this phenomenon is brought to a higher level, such that the individual entities that comprise the texture are either imperceptible or perceptually insignificant. This is not to say that total integration or fusion is required for sound mass perception, however. Indeed, total integration *cannot* be a property of sound mass, because sound mass requires an awareness of multiple sound components or sources. Were it to become so fully integrated that the impression of multiplicity were lost, then there would be nothing to distinguish a sound mass from any other auditory stream.

Expanding on Bregman's work, David Huron (2001) has demonstrated the direct application of psychoacoustic principles, including those of auditory scene analysis, in the traditional rules of voice leading in tonal counterpoint. For instance, Huron posits that the tendency for closely spaced sounds in the low range of human hearing to interact with one another creating increased sensory dissonance explains the tendency for tonal composers to voice chords with wider intervals on the bottom ("minimum masking principle"), whereas the difficulty listeners experience tracking more than three simultaneous streams explains polyphonic composers' tendency to have no more than three voices active at any given time ("principle of limited density"). We believe that these very same principles are at work in sound mass music, but with opposite aesthetic goals:

whereas tonal counterpoint preserves the individual identities of each musical line in a complex texture, sound mass music obliterates them. Thus, rather than widening intervals in the low register to minimize masking, sound mass composers use dense pitch constructions in all registers; rather than restricting the number of simultaneous voices to three or less, sound mass composers often use many more; rather than minimizing the overlapping of voices, sound mass composers overlap voices routinely; and so forth.

We further believe that the term *sound mass* encompasses a host of cross-domain mappings and conceptual blends (Nussbaum, 2007; Zbikowski, 2002) that selectively relate psychoacoustic properties of different types of sound mass compositions to various shades of meaning of the word *mass*. Sound mass has been created with many different musical resources, including instrumental, vocal, electroacoustic, and mixed media, and also many different musical techniques, including granular synthesis, tone clusters, musique concrète, musique concrète instrumentale, spectral fusion, registral extremity, rhythmic complexity, rhythmic stasis, micro-polyphony, and aleatoric and stochastic procedures. The semantic dimensions of sound masses produced in these different ways may vary considerably, ranging from literal source-bonding (in which the source-cause of a sound is clearly identified; Smalley, 1993) to mimetic suggestion to the purely abstract.

While the idea of sound mass is fairly intuitive, the many connotations of the everyday term *mass* and the variety of poetic and semantic emphases of composers and theorists have resulted in a diffuse and polysemic concept in need of clarification. The term's many senses, including formlessness (amorphous mass), materiality (physical mass), spatiality (voluminous mass), solidity (impenetrable mass), numerousness (the masses), enormity (massive), and aggregation (amassed), are not necessarily synonymous, nor is it always clear which connotations authors intend to imply and/or exclude or how those connotations relate to the experience of the listener. Indeed, while there are many published sources detailing sound mass composers' methods and aesthetic intentions, we found no sources that deal with sound mass perception empirically from the listener's perspective.

The sonic properties of a mass, the semantic associations of its *massness*, and the perceptual relations between the two may all vary considerably. For example, the rhythmically dense and ametrical pizzicati in mm. 52-29 of Xenakis' *Pithoprakta* (1956) may be perceived as a mass because of an impression of numerousness, while the opening chord of Ligeti's *Atmosphères* (1961)

may be perceived as a mass because of an impression of spatiality. These interpretations would be consistent with descriptions offered by the composers. Xenakis (1971) described his compositional method in terms of “the law of large numbers” which “implies an asymptotic evolution towards a stable state, towards a kind of goal, of *stochos*, whence comes the adjective ‘stochastic’” (p. 4), whereas for Ligeti (1983) “spatial associations play a major role in music, but the space is purely imaginary. . . [I] had been trying to suggest space, or to generate space by association in my works” (p. 92–93). Accordingly, we believe that a catalogue of mass types would be more appropriate than a singular umbrella term. However, before such a catalogue can be developed, a clearer understanding of just what constitutes a sound mass perceptually must be developed.

This problem presents several challenges. Because the term currently lacks a standard definition, it is difficult to instruct subjects to self-report their sound mass perception. Some of the parameters involved are covariant, such as timbre, register, and pitch, so threshold values may be relative rather than absolute. Depending on the compositional method, sound mass may be achieved in many different ways and may vary widely from piece to piece, employing, for instance, extremely rapid attack rate, rhythmic indeterminacy or randomness, the absence of hierarchical structuring, the absence of distinct patterns, the overabundant superposition of patterns, homogeneity of sound events, sheer numerousness of sound events, pitch saturation over a given range, similar motion among many parts, and so forth. Because of this diversity it is not possible to provide a general account of sound mass perception on the basis of any one piece, composer, or method: a series of experiments exploring the many ways sound mass has been achieved in the repertoire is required. However, it is difficult to isolate variables experimentally using real music, because there are few pieces in the sound mass repertoire in which any one variable is held constant.

However, such pieces do exist. One example is Ligeti's *Continuum* (1968, published 1970), in which the attack rate is constant from beginning to end. There are no tempo changes and the notated rhythm appears as a steady stream of eighth notes instructed to be played “as fast as possible, so that the individual tones can hardly be perceived, but rather merge into a continuum.” Ligeti is exploiting a perceptual phenomenon that has been common currency among composers for centuries in devices such as trills, tremoli, and drumrolls, but in this composition its role is foundational rather than ornamental. Although this piece is different in character and construction from his orchestral sound mass

compositions such as *Atmosphères* (1961) and *Lontano* (1967), Ligeti had similar perceptual aims:

I thought to myself, what about composing a piece that would be a paradoxically continuous sound, something like *Atmosphères*, but that would have to consist of innumerable thin slices of salami? A harpsichord has an easy touch; it can be played very fast, almost fast enough to reach the level of continuum, but not quite (it takes about eighteen separate sounds per second to reach the threshold where you can no longer make out individual notes and the limit set by the mechanism of the harpsichord is about fifteen to sixteen notes a second). As the string is plucked by the plectrum, apart from the tone you also hear quite a loud noise. The entire process is a series of sound impulses in rapid succession which create the impression of continuous sound. (1983, pp. 22–23)

Ligeti is not the only author to specify a frequency limit for sound mass fusion (“the level of continuum”); Truax (1988) and Roads (2003) both define the slightly faster value of 20 attacks/s as the threshold of continuum in granular synthesis, while Grisey (1982) specifies 20 Hz as the threshold separating the perception of pulsation from timbral fusion. The value of 20 events/s is known to be perceptually significant in other domains, having been used as the number of frames per second in some early films to exploit the visual phenomenon of apparent motion (Brown, 2014), and defining the fuzzy boundary between pitch perception and infrapitch in audition (Warren, 2008). Justin London (2004) identifies the longer value of ca. 100 ms (i.e., 10 attacks/s) as “the shortest interval that we can hear or perform as an element of rhythmic figure” (p. 27), implying that shorter values (faster event rates) are superrhythmic, beyond our ability to resolve temporally. However, we question whether attack rate alone is a sufficient predictor of sound mass perception. If Ligeti's claim that the threshold of continuum is 18 sounds/s were universally true, then all music that achieved that attack rate would be perceived as sound mass. Furthermore, there would be no variation in the perception of sound mass over the course of a piece such as *Continuum* in which the attack rate is constant.

But as is acknowledged by all who have written about *Continuum*, including Ligeti himself, the rhythm perceived by the listener is not reducible to this attack rate but rather arises from patterns in the notes. This piece has received considerable scholarly attention, with much interest aroused by this apparent contradiction between the music as notated and its perceptual reality. Several authors have gone as far as to describe *Continuum* as

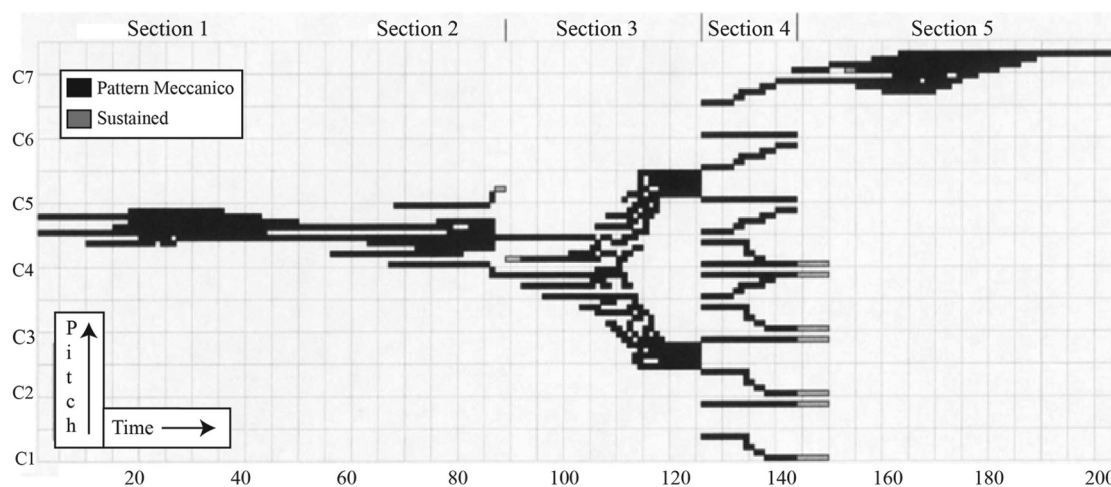


FIGURE 1. Analysis of the pitch structure in *Continuum*, © Jane Piper Clendinning 1993. Reproduced with permission.

a perceptual experiment (Peitgen, 2011), and (anachronistically) as a piece based on granular synthesis (Cambouropoulos & Tsougras, 2009). Clendinning (1993) identifies *Continuum* as one of Ligeti's "pattern-meccanico" pieces, which were inspired by his fascination with machines and mechanical sounds. The pattern-meccanico pieces are constructed from "small groups of pitches rapidly repeated in a mechanical fashion with gradual changes of pitch content," the units of which are "repeated quickly enough that the pitches almost fuse into a chord" (p. 195). Clendinning provides a helpful visualization of the varying pitch densities and compasses that unfold over the course of the piece (Figure 1). This graph, essentially a condensation of the score, depicts pitch on the y-axis and time on the x-axis. Any "division" (they are not really measures, as Ligeti's barlines are dotted, and he specifically prohibits metric accent) in which a pitch appears in a meccanico pattern is represented by a black box, and any division in which a pitch is sustained is represented by a grey box. The thick, black sections of the graph in Figure 1 thus represent areas of high pitch density, whereas areas with white space between horizontal bands represent defined intervals. If, as is commonly assumed to be the case, pitch density correlates with sound mass perception, we would thus expect the thick dark areas to correspond with the places in the piece that exhibit the highest degree of perceptual fusion.

A comprehensive analysis of the piece's rhythmic organization is provided by Petersen (2008), who describes it as composed of three types of building blocks: "Pendel, Sägezahn und Welle" (pendulum, sawtooth, and wave). The first is the directionless alternation of two notes, the second is a unidirectional pattern

of three notes or more, always ascending in the left hand (LH) and descending in the right hand (RH), and the third is a bidirectional oscillating pattern of three notes or more. The superposition of such patterns, especially when the patterns in RH and LH are of different lengths or are asynchronous, gives rise to many possible rhythmic groupings depending on the directed attention of the listener. In the excerpt analyzed in Figure 2, the left hand plays a pendulum pattern throughout, while the right hand begins with a pendulum and moves to a sawtooth. As this figure shows, the rhythm (after the right hand has initiated the sawtooth pattern) may be perceived in several different ways depending on the directed attention of the listener: as a steady stream of eighth notes if the listener attends to the most basic level of rhythmic organization (*Rhythmus* on line 3), as a two-against-three pattern if the listener attends to the groupings played by each hand (*Spielfiguren*), as an overall pulse of six eighth-notes if the listener attends to completed cycles of the two-against-three pattern (*Patterns*), as a group of four non-repeating pitch combinations followed by a repeated G – B $\flat$  dyad if the listener attends to repetition (*Zweiklang-Repetitionen*), and as a period of three eighth-notes if the listener attends to the lowest pitch F as the marker of periodicity (*Unterstimme*). In other words, the rhythm of the piece is *emergent*, determined by factors such as pitch organization and periodicity that at times can cause the music to segregate perceptually into multiple streams if they attract the listener's attention. Over the course of the piece, changing periodic patterns and pitch structures of varying degrees of complexity exploit emergent rhythm as a primary organizational feature of the piece, with



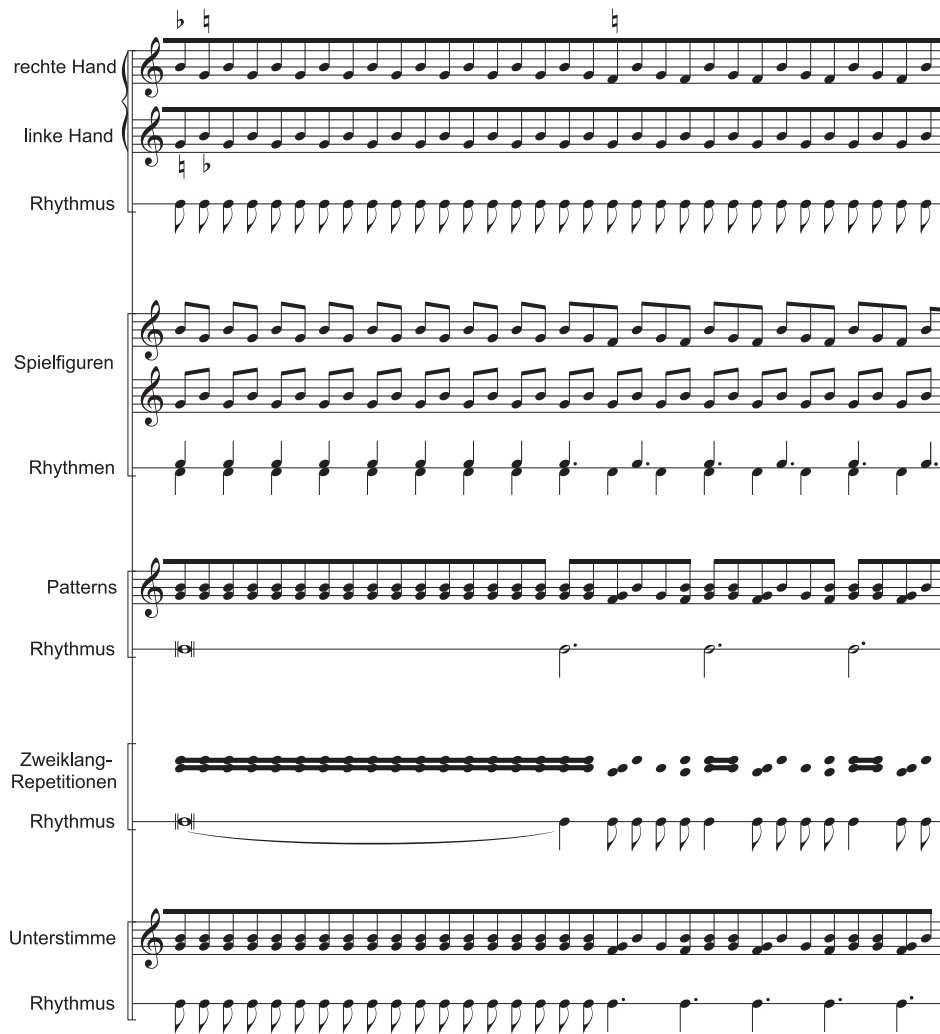


FIGURE 2. Analysis of component rhythms in *Continuum*, © Peter Petersen 2008. Reproduced with permission.

potentially significant implications for sound mass fusion which will be explored below.

An interesting consequence of the construction of *Continuum* as a series of dyads is that the greater the number of pitches present in any given pattern, the more distinct its emergent rhythm will become. Other pieces, such as Ligeti's orchestral works *Lontano* and *Atmosphères*, may use density or complexity of pitch structure as a mechanism by which to evoke sound mass, but here the situation is rather different. Because only two pitches are articulated at any given instant, emergent rhythm and pitch structure are likely to operate at cross-purposes in terms of perceptual fusion. This inverse relation is illustrated visually in Figure 3: the pendulum pattern in the first system appears as a continuous stream of identical dyads, but the wave that emerges

as more notes are added to the pattern below segregates into a series of discrete gestures with clear high and low points. At a tempo of 18 attacks/s, the pendulum would have a periodicity of 56 ms (ca. 1080 beats per minute) and would certainly be superrhythmic, whereas the longest wave patterns (eight notes long) would have a periodicity of 444 ms (ca. 135 beats per minute), well within the range London identifies for beat or pulse perception (London, 2004, p. 46). Thus, more complex pitch structures, which under other circumstances may result in greater sound mass integration, here result in emergent rhythm likely to cause auditory stream segregation. Conversely, simpler pitch structures, which might otherwise cause stream segregation, result in the attenuation or abolition of emergent rhythm and hence promote sound mass fusion.

The image displays four systems of musical notation for the piece 'Continuum'. Each system consists of two staves, one for the right hand (r.H.) and one for the left hand (l.H.). The notation is highly rhythmic, featuring a dense sequence of eighth notes with varying rests, creating a complex, layered texture. The systems are labeled with measure numbers: 50, 56, 62, and 68. The overall effect is a continuous, evolving stream of sound.

FIGURE 3. Pendulum and wave patterns of varying lengths in *Continuum*, © Peter Petersen 2008. Reproduced with permission.

Because, as we have seen, emergent rhythm and other cues likely to promote stream segregation vary with musical organization, it is likely that the degree of sound mass perception will vary over the course of the piece, being highest when such cues are most attenuated, and lowest when they are most clear. Cambouropoulos and Tsougras (2009) have analyzed the piece explicitly in terms of auditory scene analysis, presenting “a theoretical-analytical discussion of the implication of perceptual principles on our understanding on the structure of *Continuum*,” and thereby analyzing the differences between the music as “notated” (a continuous series of eighth notes) and “perceived” (superposed auditory streams of variable organization). For example, in Figure 4, two excerpts from the piece are depicted first in “piano roll” notation, then as they appear in the score, and finally as they are “perceived,” with each staff representing a distinct auditory stream.

However, Cambouropoulos and Tsougras (2009) did not test their findings empirically, relying instead on

musical analysis to determine how the piece will be perceived and stating that “in future work, a perceptual study may be carried out testing the hypothesis of streaming in this musical piece” (p. 127). This is one of the aims of the present study.

*Continuum* is best known in its original version for harpsichord, but Ligeti also produced two other versions: one for two player pianos, the other for barrel organ. The piece thus provides an interesting opportunity to compare the effect of different timbres on sound mass perception without artificially modifying the composer’s work. As can be seen in the graphs on the right in Figure 5, the harpsichord’s attack is much sharper and decay much more rapid than those of either the piano or the organ. As seen in the graphs on the left, the harpsichord also has a considerably greater spectral extent leading to a higher spectral centroid, which can be expected to correlate with a higher degree of “brightness” and timbral saliency. These perceptual cues would seem likely to draw attention to the individual identities

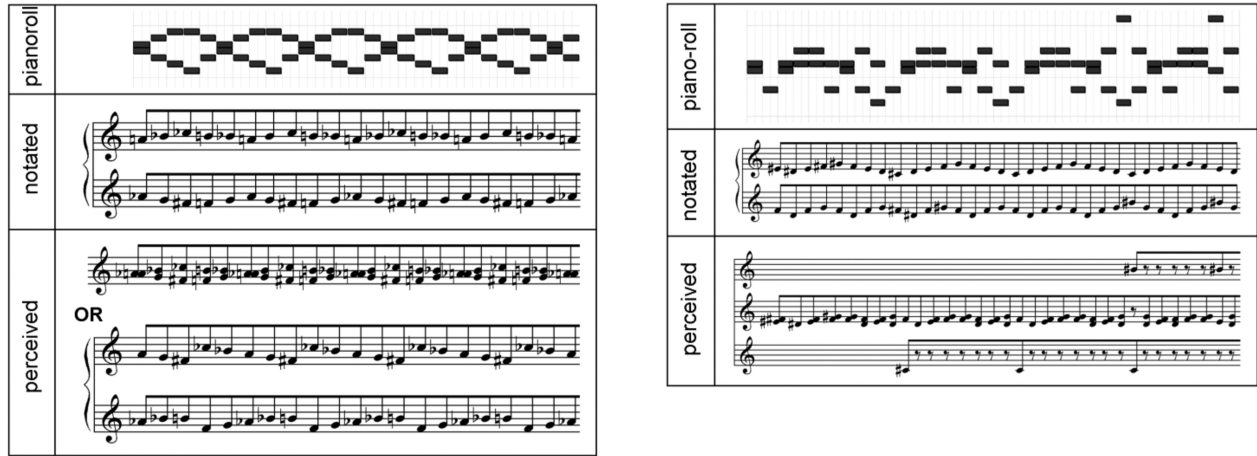


FIGURE 4. Piano-roll notation, notation in the score, and putative perceived organization as interpreted by Cambouropoulos and Tsougras (2009). © 2009 by *Journal of Interdisciplinary Music Studies*. Reproduced with permission.

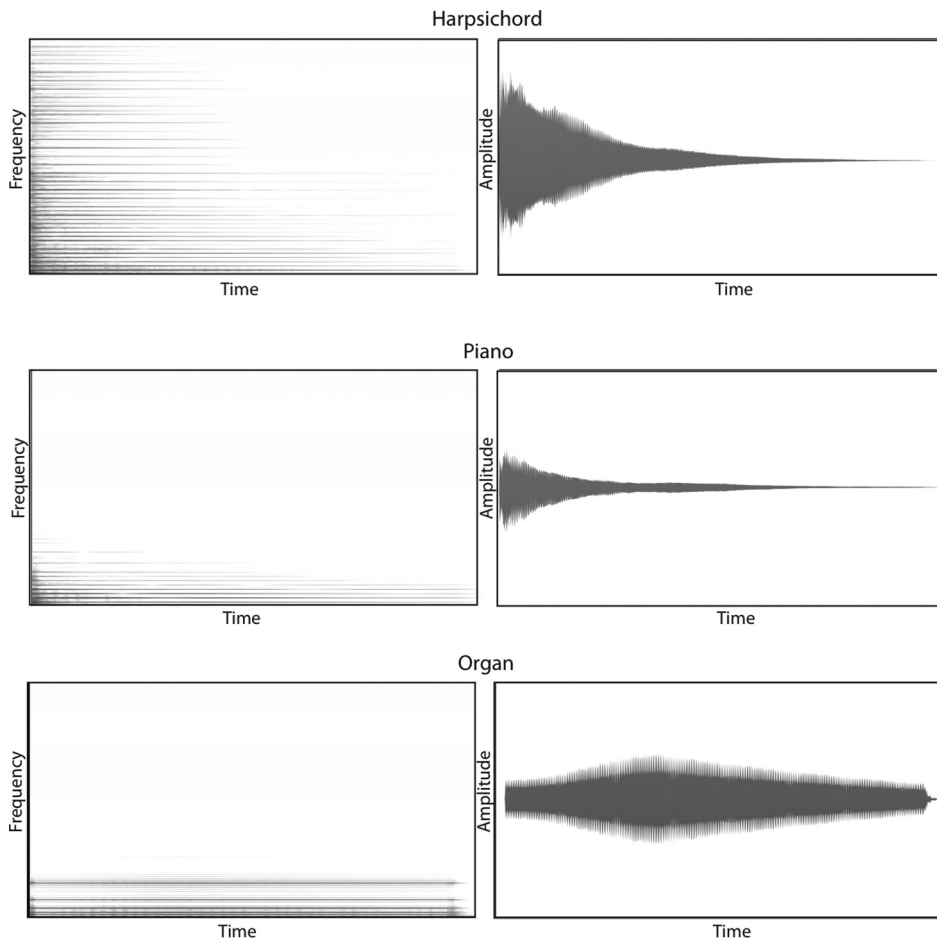


FIGURE 5. Spectrographs (left panels) and amplitude envelopes (right panels) of the pitch G4 played on the harpsichord (A), the piano (B), and the organ (C).

of harpsichord notes, creating obstacles to perceptual fusion. The organ, on the other hand, has a very soft attack, homogeneous sustain, and much lower spectral centroid, which would seem to provide fewer cues for the demarcation of individual notes and therefore to promote fusion. All other things being equal, then, we would expect the harpsichord to yield an overall lower degree of sound mass perception than the other two instruments, or to require a higher attack rate in order to achieve the same degree.

To summarize, we sought to experimentally evaluate the following hypotheses with respect to the three versions of Ligeti's *Continuum*:

1. The degree of sound mass perception will vary over the course of the piece according to musical organization.
2. The rate of attack required for sound mass perception is not an absolute value as has been previously suggested. It will interact with other musical and psychoacoustic factors such as timbre, register, and musical organization.
3. Pitch density correlates positively with sound mass perception but this correlation is not one-to-one as is often assumed. Pitch density interacts with other musical and psychoacoustic factors in sound mass perception. In *Continuum*, the interaction between pitch structure and emergent rhythm will be especially consequential.
4. The three instrumental timbres will differ in their proclivity for sound mass perception in the following increasing order: harpsichord, piano, organ. This ranking is based on sharpness of attack, brightness of timbre (spectral centroid), and rate of decay.

### Experiment 1

We sought to determine whether sound mass perception in *Continuum* was constant (like its attack rate) or variable (like its musical organization). We also sought to determine whether or not instrumental timbre affected sound mass perception when attack rate and musical organization remained constant. By comparing listeners' self-reported continuous responses of degree of sound mass perception between the three versions of *Continuum*, we expected to observe: (1) changing responses over the course of the piece, varying according to register, pitch organization, and emergent rhythm, and (2) overall lower ratings for the harpsichord, medium for the piano, and higher for the organ based on their acoustic characteristics.

### METHOD

*Participants.* Forty participants, between 18 and 64 years of age ( $M = 25$ ,  $SD = 9.6$ ), completed the experiment. Twenty-one participants reported formal music training, ranging from 5 to 52 years of practice ( $M = 13.0$ ,  $SD = 9.8$ ). The remaining participants reported no music training at a collegiate level and no more than a year of formal music training during childhood ( $M = 0.35$ ,  $SD = 0.49$ ). Prior to completing the experiment, all participants passed a pure-tone audiometric test using a MAICO MA 39 (MAICO Diagnostic GmbH, Berlin, Germany) audiometer at octave-spaced frequencies from 125 Hz to 8 kHz (ISO 389-8, 2004; Martin & Champlin, 2000) and were required to have thresholds at or below 20 dB HL to proceed. All participants completed the same task and were paid \$5 CDN as compensation. They all signed informed consent forms prior to participating in the experiment.

*Stimuli.* The stimuli consisted of recordings of the three versions of *Continuum*: for harpsichord, piano, and organ. Besides instrumentation, all musical parameters were identical in the three versions, following Ligeti's score exactly. The stimuli were created in Finale 2011 rendered into sound with the Garritan Sound Library and were 3 min 48 s in duration with a consistent tempo of 16 attacks/s. Audio signals were sampled at 44.1 kHz with 16-bit amplitude resolution. Additionally, a recording of *Coulée* (Ligeti, 1997), a piece for organ, was used as a practice trial to familiarize participants with the type of music they would be hearing and the experimental interface.

*Procedure.* Participants completed the experiment individually inside an Industrial Acoustics model 120-act3 double-walled sound isolation booth (IAC Acoustics, Bronx, NY). The pieces were amplified with a Grace Design m904 monitor system and heard over circumaural Sennheiser HD280 Pro earphones (Sennheiser Electronic GmbH, Wedemark, Germany) at an average level of 60 dB SPL for all participants. The different instrument versions were presented once each in a randomized order for each participant.

The participants completed the experiment on an iPad interface (Apple Computer, Inc., Cupertino, CA). The interface was created in the mrmr OSC controller application (Redlinger, 2008) and consisted of four play buttons, one for the practice trial and one for each of the three test trials, as well as a continuous slider scaled from "No Sound Mass" to "One Complete Sound Mass." The iPad communicated via OpenSoundControl (Center for New Music and Audio Technologies, Berkeley, CA) messages over a wifi network with a Max/MSP version 5.1.9 (Cycling '74, San Francisco, CA) patch run



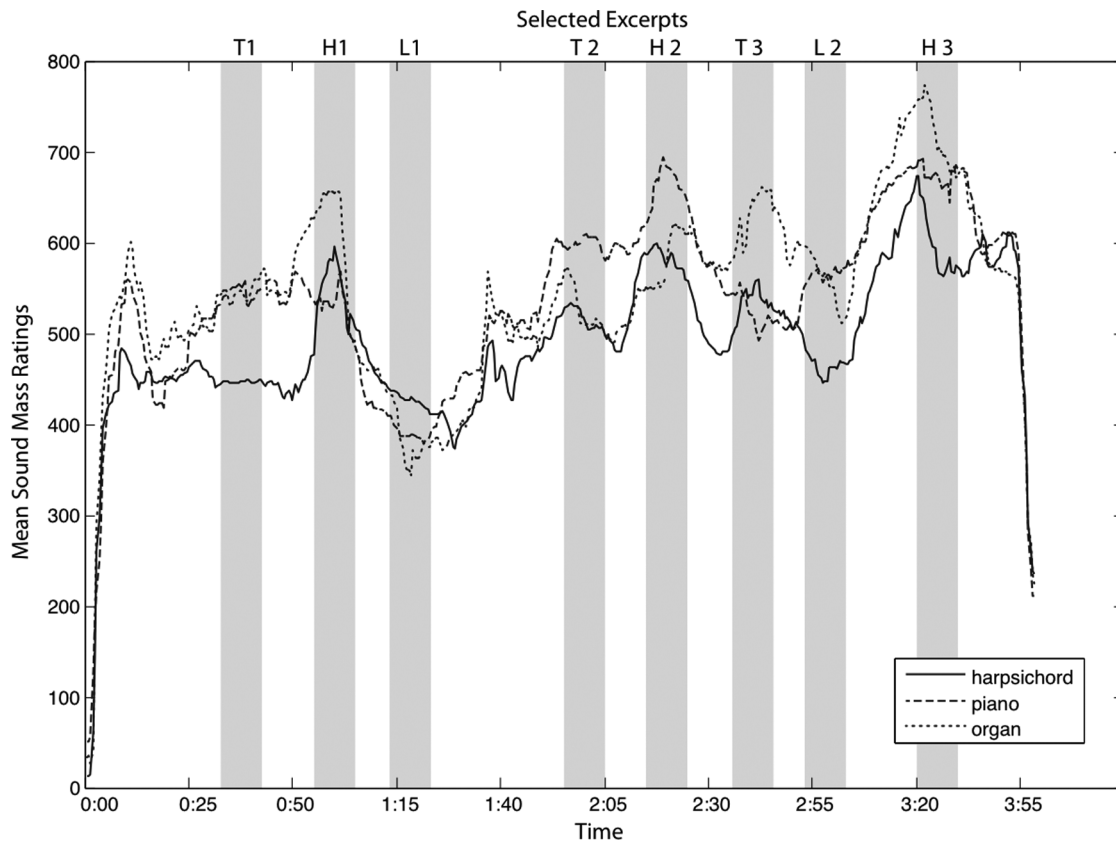


FIGURE 6. Average sound mass rating profiles. Selected excerpts are shown by vertical grey bars and are labeled at the top.

on a Mac Pro computer (Apple Computer, Inc., Cupertino, CA). The Max/MSP patch was designed to randomize and play the stimuli as well as record and output the ratings.

Participants continuously rated sound mass perception while listening to each version of *Continuum*. Data were collected at a rate of 2 Hz. In the instructions, they were given the following description:

Sound mass is a type of music in which sounds lose their individual identities and are heard as a totality; it is similar to the way that voices in a crowd can become difficult to distinguish as individuals, but become integrated into the sound of the crowd as a whole.

Participants were informed that a rating of “No Sound Mass” meant that the sounds they heard retained their distinct, individual character. A rating of “One Complete Sound Mass” meant the sounds they heard were completely integrated into a single mass. Finally, an intermediate rating would represent situations in which some but not all of the sounds were integrated

into a sound mass, or in which multiple simultaneous masses were perceived, or in which sounds were partially integrated but still retained distinct individual identities. In the results section we will refer to ratings towards “One Complete Sound Mass” as high ratings and ratings towards “No Sound Mass” as low ratings.

#### RESULTS

The averages of the continuous ratings over all participants for each instrumental version are shown in Figure 6. Based on this visual representation of the results, we chose eight excerpts to statistically analyze. We specifically chose excerpts that represented areas of high ratings and low ratings so we could musically analyze these sections and use the theoretical analysis to explain the differences in the perceptual ratings. Additionally, we wanted to examine excerpts where there was a visual difference between the different instrumental pieces so that we could search for a timbral explanation for varied ratings between the instruments. All selected excerpts were 10 s in duration and therefore consisted of 20 data

points each. Of the eight excerpts, three represented areas of high sound mass ratings (labeled H1, H2, H3 in Figure 6) and two represented areas of low sound mass ratings (L1, L2). Additionally, three excerpts (T1, T2, T3) were chosen for comparisons of the three different instrumental timbres. In all analyses of variance, Mauchly's test was used to test for violations of sphericity, and if significant, the degrees of freedom were corrected using the Greenhouse-Geisser epsilon if  $\epsilon \geq .75$  and the Huynh-Feldt epsilon otherwise. Original degrees of freedom, epsilon, and the corrected  $p$ -value are reported where appropriate.

An initial mixed analysis of variance (ANOVA) was conducted primarily to determine whether the between-subjects factor of musicianship (musician, nonmusician) interacted with any of the within-subject factors of instrument (harpichord, piano, organ) and excerpt (H1, H2, H3, L1, L2, T1, T2, T3). The main effect of musicianship was not significant, nor were any of its interactions with the other factors,  $F < 1.20$ ,  $p \geq .35$  in all cases. There was a marginally significant main effect of instrument,  $F(2, 76) = 2.58$ ,  $p = .08$ ,  $\eta_p^2 = .06$ , and a significant main effect of excerpt,  $F(7, 266) = 10.50$ ,  $\epsilon = .51$ ,  $p < .01$ ,  $\eta_p^2 = .22$ . More importantly for our purposes, the interaction between instrument and excerpt was also significant,  $F(14, 532) = 3.90$ ,  $\epsilon = .86$ ,  $p < .001$ ,  $\eta_p^2 = .09$ . Paired-sample  $t$ -tests comparing the mean across the highly rated excerpts against that of the lowly rated excerpts (H1, H2, H3 vs. L1, L2) were statistically significant for each instrument,  $t(39) \geq 4.36$ ,  $p < .001$ . We can conclude that sound mass perception varies over the course of a piece depending on musical attributes and is affected by the instrument playing the piece, but music training does not influence sound mass ratings.

To further evaluate differences among ratings of the three instrumental versions, we ran a repeated-measures ANOVA, examining the within-subject factors of excerpt and instrument on excerpts T1, T2, T3, and H3, specifically because these were the excerpts with a visual difference between the ratings of the different instrumental versions. We did not include musicianship as a factor, because we had found no differences between the groups in the previous analysis. Figure 7 is a plot of the mean rating of each instrumental version for excerpts T1, T2, T3, and H3. Generally the ratings increase across these excerpts, with the piano and organ ratings being typically higher than the harpsichord ratings. However, the piano ratings become lower than the other instrumental versions in excerpt T3, and this difference results in a significant interaction between instrumental version and excerpt,  $F(6, 234) = 4.01$ ,  $p = .001$ ,  $\eta_p^2 = .093$ . We conducted 12 paired-sample

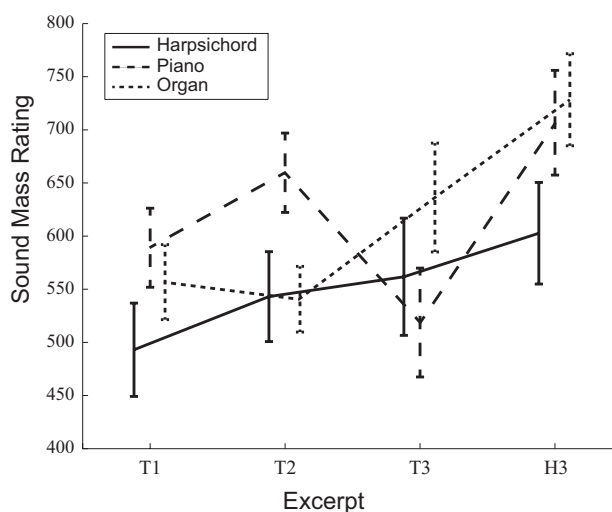


FIGURE 7. Average sound mass ratings for excerpts with apparent instrumental differences.

$t$ -tests between the three pairs of instruments for each of the four excerpts. After Bonferroni-Holm correction, only one significant difference was found: the organ means are higher than the harpsichord means in excerpt H3,  $t(39) = -3.33$ ,  $p(\text{corrected}) = .036$ . We conclude that, overall, the instrumental differences are not reliable throughout the piece.

#### DISCUSSION

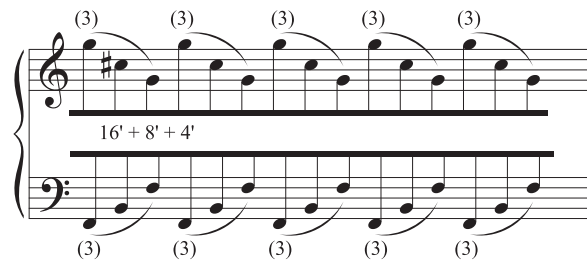
The observed results are coherent with the organizational properties of the music. The three highly rated excerpts had very short, synchronized periodicities. The first (Figure 8A) consists of two notes separated by a major second (F#4 – G#4), repeated with the fastest possible periodicity of one note ( $0.063\text{s} = 16$  cycles/s). The second (Figure 8B) consists of a 6-note whole-tone chord spread over more than three octaves (F2-B2-F3 – G4-C#5-G5), repeated with a very fast periodicity of three notes ( $0.19\text{s} \approx 5$  cycles/s). Technically the compass of this second excerpt is even wider, because the 16' and 4' couplers double each note an octave below and an octave above, resulting in a spread of over 5 octaves. Therefore the two excerpts are virtually opposite in their pitch density, the first consisting of a single small interval in a very narrow compass, the second of a chord built of larger intervals spread over a very wide compass. Excerpt H3 (Figure 8C) is a rapid minor-second trill (Fb6 – Eb6), with a periodicity of two notes ( $0.125\text{s} = 8$  cycles/s). The commonality appears not to be pitch density, but rapid, synchronous, periodic repetition.

The two sections with low sound mass ratings (Figure 9) also exhibit striking organizational similarity. Both consist

A. Excerpt H1



B. Excerpt H2



C. Excerpt H3

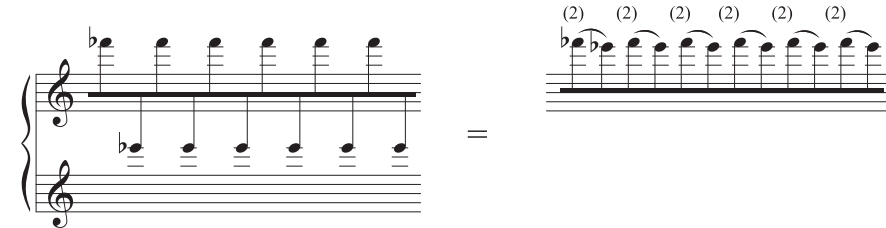
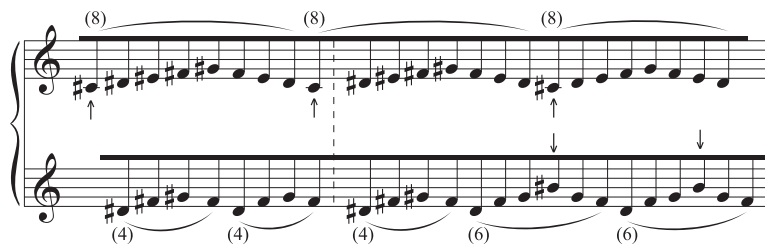


FIGURE 8. The three excerpts (H1, H2, and H3) with high average sound mass ratings.

A. Excerpt L1



B. Excerpt L2

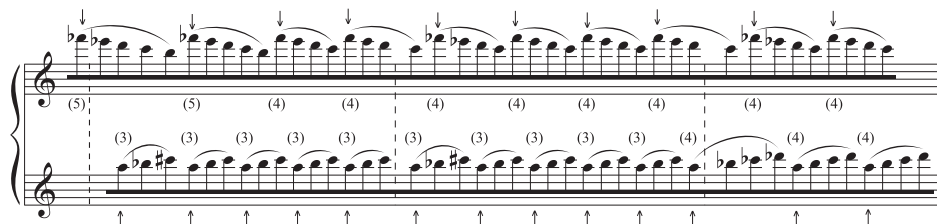


FIGURE 9. The two l excerpts (L1 and L2) with low average sound mass ratings.

of patterns between the two hands that are either unequal or asynchronous. The periods individually tend to be relatively longer, and the highest and lowest pitches are placed in positions of contour and/or register that draw attention to them, potentially causing them to segregate into separate streams.

It would thus appear that regular patterns with short periodicities and synchronous (or singular) onsets promote sound mass fusion, while patterns with unequal periodicities or asynchronous onsets tend to promote segregation. Even though the attack rate was constant at 16 notes/s, the emergent rhythm caused by pitch organization resulted in the perception of periods that could be much longer, disintegrating the perception of fusion. This tendency would appear to override pitch density as a predictor of sound mass fusion, as is evident in a comparison between a chromatically saturated excerpt like L2 with a widely spaced excerpt like H2.

A simple explanation for any relations that may exist between instrumental timbre and sound mass fusion is not as forthcoming. The three sections that were selected for apparent differences in ranking for instrumental timbres exhibit little similarity in terms of musical organization: T1 is a chromatic section with unequal periods similar to L2 but in a middle register (F#4 – Bb4), T2 is a volatile section with constantly changing periodicity and diverging registers between the hands, and T3 is a major second trill (B5 – C#6) similar to H3 but with sustained notes in the left hand. In spite of the visual appearance of the graph, the distribution of the data for the three instrumental versions in these sections overlap a great deal. Although the tendency of the harpsichord version to have lower mean ratings compared to the other two lends some support to our intuitive assumption that timbre and perceptual fusion are related, the effect is not statistically reliable, and further research will be needed to draw any strong conclusions about the nature of their relation.

## Experiment 2

Having found in Experiment 1 that sound mass perception in *Continuum* varied significantly over the course of the piece in accordance with changing musical organization, we wondered whether varying selected musical parameters would affect these results. We selected three excerpts from the piece, two with high average ratings and one with low average ratings of sound mass perception (labeled H1, L1, and H2 above), and created multiple versions of those sections by playing them at different speeds and transposing them to different octaves. We expected to observe that (1) ratings would

increase when excerpts were transposed to very high or very low registers since human hearing is most acute in the middle register or “spectral dominance region” (Huron, 2001), (2) sound mass perception would be positively correlated with attack rate, and (3) instrumental timbre would affect ratings on shorter excerpts.

## METHOD

*Participants.* The 40 participants (23 females) were between 18 and 66 years of age ( $M = 25.0$ ,  $SD = 9.6$ ). Twenty participants reported formal music training ranging from 8 to 19 years of practice ( $M = 12.0$ ,  $SD = 3.2$ ). The remaining participants reported no music training at a collegiate level and less than two years of formal music training during childhood ( $M = 0.55$ ,  $SD = 0.83$ ). Prior to completing the experimental task, participants passed the hearing test detailed in the previous experiment. All participants completed the same task and were paid \$10 CDN as compensation. They all signed informed consent forms prior to participating in the experiment.

*Stimuli.* As noted above, excerpt H1 (beginning at division 49 in the score, Figure 8A) was rated very highly for sound mass perception and consists of a repeating dyad (F#4 – G#4) with a periodicity of 1 (i.e., pattern repeated each note). Excerpt L1 (beginning at division 68, Figure 9A) had a very low rating and consists of a complex superimposition of several patterns of differing periodicities and pitch organizations. Excerpt H2 (beginning at division 126, Figure 8B) was rated very highly and consists of an arpeggiated six-note chord with a periodicity of 3 (LH: F2-B2-F3; RH: G5-C#5-G4) with each note doubled at the octave above and the octave below.

Seventy-five versions of each excerpt were created in Finale 2011 with the Garritan Sound Library, using three instrumental timbres (harpsichord, piano, organ), five speeds (8, 12, 16, 20, and 24 attacks/s), and five transpositions (original pitch,  $\pm 1$  and  $\pm 2$  octaves). Audio signals were sampled at 44.1 kHz with 16-bit amplitude resolution. Loudness-equalized stimuli were presented via the same equipment outlined in Experiment 1 at an average level of 60 dB SPL.

Ten participants who did not participate in Experiment 2 completed a loudness equalization task in the PsiExp software environment (Smith, 1995). The task was completed in three blocks, one for each excerpt, presented in random order. Each block consisted of 75 samples from the three instruments, five octaves, and five attack rates. The reference samples for each block were the organ versions of each excerpt played at the original octave at 16 attacks/s, and the remaining 44

TABLE 1. Linear Mixed Effects Model with Type III Wald F Tests for Perceived Sound Mass Ratings.

|   |               | H1<br>( $R^2 = .42$ ) | H2<br>( $R^2 = .44$ ) | L1<br>( $R^2 = .42$ ) |
|---|---------------|-----------------------|-----------------------|-----------------------|
|   | df            | F                     | F                     | F                     |
| Music Training  | 1, 2587-2700  | 0.26                  | 0.84                  | 0.34                  |
| Instrument  | 2, 2587-2700  | 0.00                  | <b>176.41***</b>      | <b>14.34***</b>       |
| Octave Change   | 4, 2587-2700  | 0.32                  | <b>28.60***</b>       | <b>51.44***</b>       |
| Attack Rate   | 4, 2587-2700  | 0.27                  | <b>68.39***</b>       | <b>76.38***</b>       |
| Music Training * Instrument                               | 2, 2587-2700  | <b>3.65*</b>          | <b>12.46***</b>       | 2.47                  |
| Music Training * Octave Change                            | 4, 2587-2700  | <b>5.22***</b>        | 0.55                  | <b>7.80***</b>        |
| Instrument * Octave Change                                | 8, 2587-2700  | 0.01                  | <b>5.33***</b>        | 1.81                  |
| Music Training * Attack Rate                              | 4, 2587-2700  | 1.58                  | 1.74                  | 1.00                  |
| Instrument * Attack Rate                                  | 8, 2587-2700  | 0.05                  | 1.93                  | <b>2.38*</b>          |
| Octave Change * Attack Rate                               | 16, 2587-2700 | 0.02                  | 0.77                  | <b>2.11**</b>         |
| Music Training * Instrument * Octave Change               | 8, 2587-2700  | 1.07                  | 1.48                  | 0.73                  |
| Music Training * Instrument * Attack Rate                 | 8, 2587-2700  | 0.69                  | 0.75                  | 0.32                  |
| Music Training * Octave Change * Attack Rate              | 16, 2587-2700 | 0.62                  | 1.17                  | <b>1.77*</b>          |
| Instrument * Octave Change * Attack Rate                  | 32, 2587-2700 | 0.01                  | 0.99                  | 1.18                  |
| Music Training * Instrument * Octave Change * Attack Rate | 32, 2587-2700 | 0.92                  | 0.80                  | 0.43                  |

Notes:  $N = 2850$ . All predictors are sum coded factor variables. Results of Type III Wald F test: \* $p < .05$  \*\* $p < .01$  \*\*\* $p < .001$ . The following random effects were included: random intercepts for participants and stimuli.

samples in each block were presented individually in a randomized order. The level of each one was adjusted to equal the loudness of the reference. The median adjustment rating from the 10 participants' ratings for each sample was then used to determine its level in Experiment 2. Only samples with an attack rate of 8, 16, and 24 attacks/s were equalized by participants and then, based on those results, an appropriate loudness level for each sample with an attack rate of 12 or 20 attacks/s was linearly interpolated between the level adjustments of the surrounding attack rates.

*Procedure.* The equipment used was the same as the first experiment, although in this experiment the iPad interface was designed with TouchOSC (Hexler.net, Berlin, Germany) which allowed for OpenSoundControl messages to be sent not only from the iPad to the MaxPatch, but also from the MaxPatch back to the iPad. This was necessary to reset the slider to the start position after each trial, ensuring independent ratings across trials. The interface consisted of a continuous scale with Likert-type labels from 1 to 9. The anchor corresponding to 1 was labeled "No Sound Mass" while the anchor corresponding to 9 was labeled "1 Complete Sound Mass." Additionally, there was a button labeled "Next" in the top right corner.

Participants were given the same definition of sound mass as the previous experiment and instructed to complete a single rating per trial after hearing the excerpt. Upon completing the rating, they were cued to press the next button, which would reset the scale and play the next sample.

## RESULTS

Statistical analyses were completed with a linear mixed model method (West, Welch & Galecki, 2007), which performs a regression-like analysis while controlling for random variance in factors such as participant and stimuli. Because each participant rated all of the stimuli, the model included crossed random effects for subject and item (Baayen, Davidson & Bates, 2008). Analyses were completed with the software R (3.0.2) using the lmer function from the lme4 package (Bates, Maechler, & Bolker, 2011) and the pamer.fnz function from the LMERConvenienceFunctions package (Tremblay & Ransijn, 2013). To decrease a skew in the ratings, we performed a logit transform on the data, which resulted in a more normal distribution ranging from about  $-6$  to  $6$ .

We created a model for each excerpt to examine music training, instrument, octave, and attack rate as factors influencing sound mass perception. The results from a Type III F test for each excerpt are displayed in Table 1. Consistently with the previous experiment, the influence of the examined factors differed across excerpts, confirming that the musical structure influences the perception of sound mass. Furthermore, music training alone was not a significant factor for any of the three excerpts; however interactions involving this factor were sometimes significant.

As shown in Figure 10, significant differences were observed among the ratings for the three instruments in this experiment. The excerpts that were highly rated in Experiment 1 (H1 and H2) were once again more highly rated overall in Experiment 2. However, the ratings for



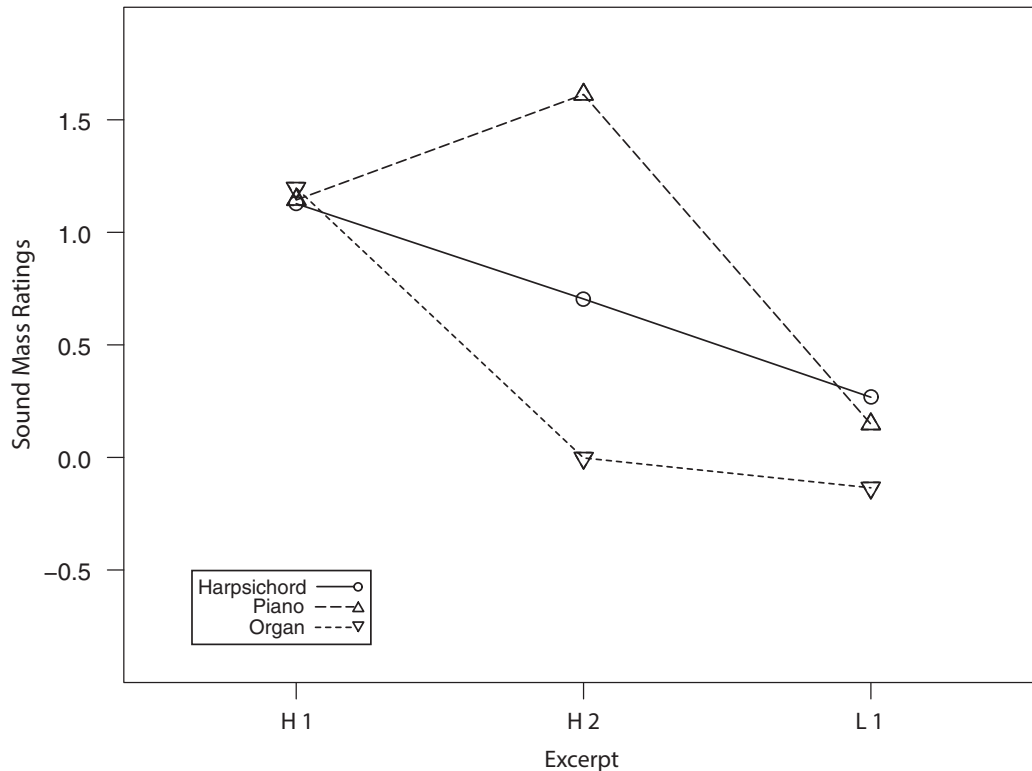


FIGURE 10. Overall sound mass ratings for three musical excerpts played with different instruments.

H1 (Figure 8A), consisting of a single major-second dyad, are much more consistent between instruments than H2 (Figure 8B), which consists of major seconds and tritones spread over a much wider compass. L2 (Figure 9B), which, as discussed above, contains several perceptual cues likely to promote stream segregation, was rated lower overall and with greater consistency between instruments than was H2.

Figure 11 represents the mean sound mass ratings across the different attack rates for each octave transposition. We originally predicted that extreme registers, both high and low, would have high sound mass ratings, because human hearing is less discriminating in those registers. However, stimuli in the lowest octaves had consistently high ratings and stimuli in the highest octave were rated lower, despite changes in other variables; so although we were correct to think that register would be a significant factor, the tendency is a monotonically decreasing pattern rather than the parabolic one we had expected. A similar linear tendency was observed for speed: increased attack rate was consistently correlated with increased sound mass ratings. We also originally predicted that a faster attack rate would be required for sound mass perception in the

middle register, and relatively slower rates would be sufficient in the registral extremes. However, there was little interaction between attack rate and octave.

In excerpts H1 and H2, significant interactions between music training and instrumental timbre with respect to sound mass ratings were observed (Figure 12). In H1, musicians rated harpsichord and piano stimuli slightly lower than did nonmusicians, but rated organ stimuli higher than did nonmusicians. In H2, nonmusicians' ratings for harpsichord stimuli were higher than musicians' ratings, but both groups rated organ stimuli very similarly. In L1, the musicians' harpsichord stimuli ratings are lower than were the nonmusician's ratings — a trend that was apparent in all three excerpts — but the difference in this case was not statistically significant.

A significant interaction between music training and octave transposition was observed in H1 and L1 (Figure 13). Both groups typically gave higher sound mass ratings in lower octaves and lower ratings in higher octaves, but the effect at both ends was more potent in nonmusicians. In H2 nonmusicians gave higher ratings at all transpositions, but no significant interactions were observed.

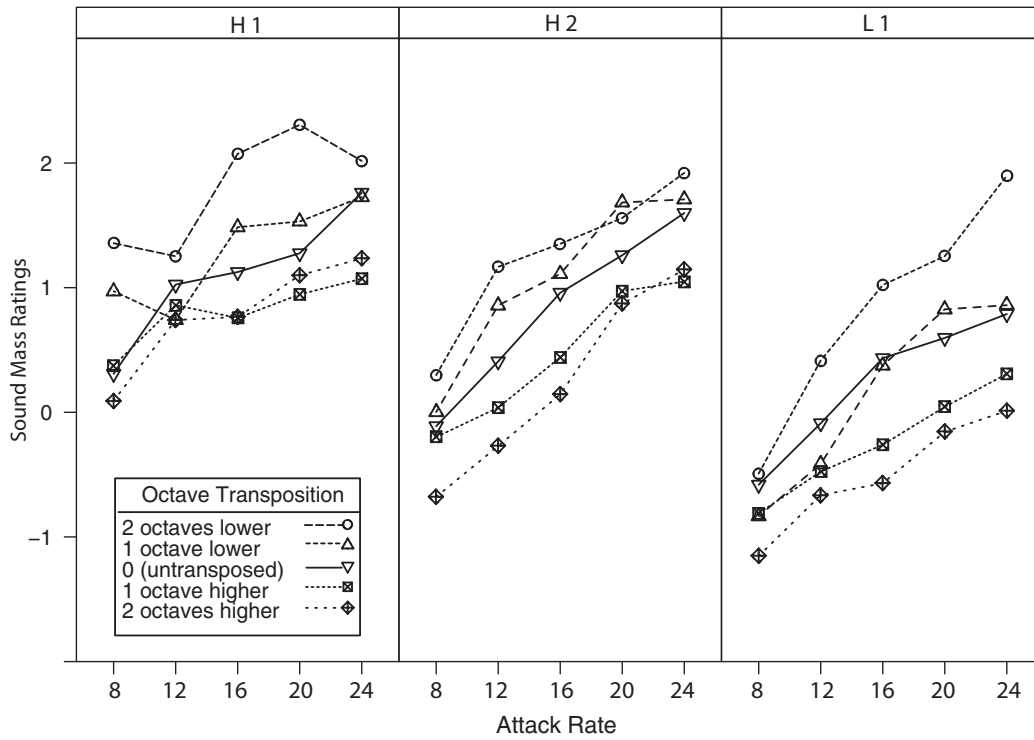


FIGURE 11. Sound mass ratings at different attack rates and octave transpositions.

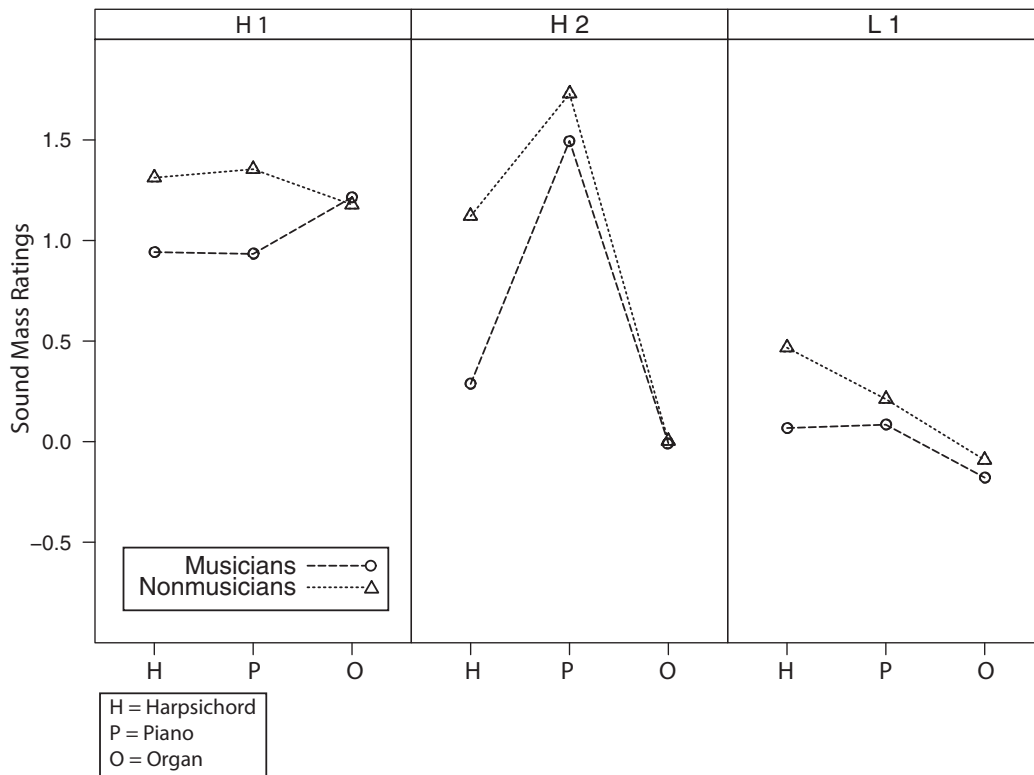


FIGURE 12. Sound mass ratings for different instruments by musicians and nonmusicians.

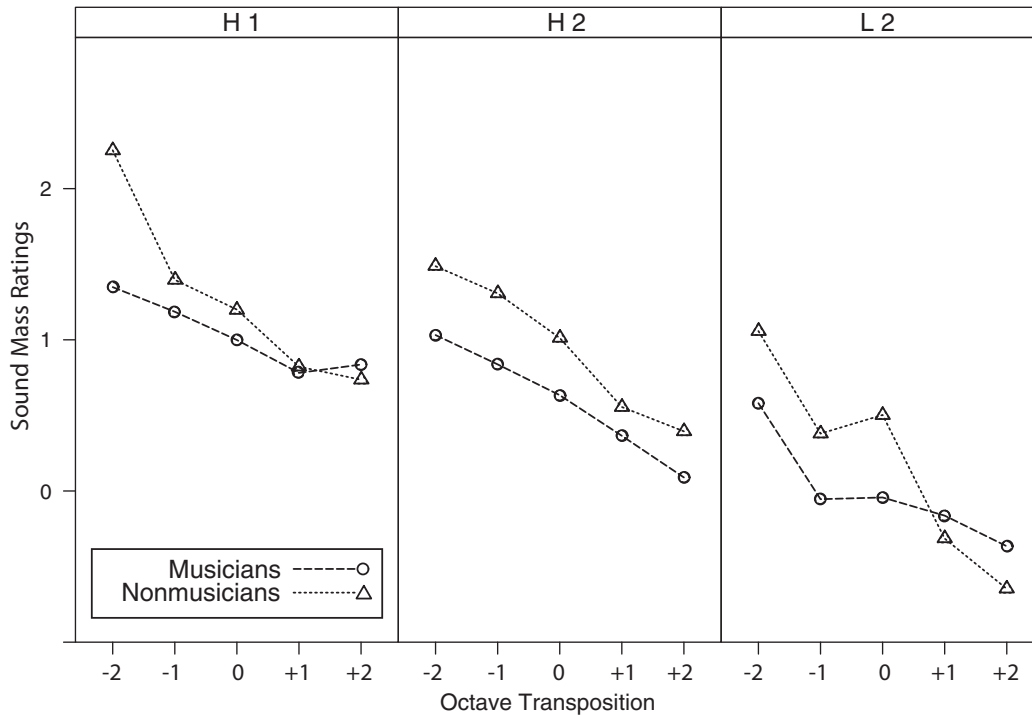


FIGURE 13. Sound mass ratings for different octave transpositions by musicians and nonmusicians.

#### DISCUSSION

The evidence we observed suggests that Truax, Roads, and Ligeti are right to claim that faster attack rate (or granulation) results on average in a higher degree of continuous perceptual fusion, which we are here calling sound mass. However, we have not found evidence supporting an absolute threshold attack rate, at least not within the range of speeds we tested (up to 24 attacks/s). Other factors such as register, timbre, emergent rhythm, pitch structure, and subjective context appear to be relevant as well, as the principles of auditory scene analysis led us to predict.

As shown in Figure 10, sound mass ratings for excerpt H2, whose pitch density is lower and whose compass is far greater than the other two, varied much more across the different instruments than did the ratings for the other excerpts. Excerpt L2, whose unequal periodicities and “outlier” pitches at the crest and nadir of wave-like pitch patterns are likely to promote stream segregation, had overall lower ratings in Experiment 2, as it was the case in Experiment 1.

Figure 11 shows that sound mass ratings for all three excerpts followed an approximately linear pattern of positive association with attack rate (higher rate = higher rating), and a negative association with octave transposition (lower octave = higher rating). However,

the start and end points of these linear trajectories varied considerably among the three musical excerpts, calling into question the notion of an absolute value in either parameter that would be universally applicable regardless of musical content or context.

The fact that sound mass ratings can be affected significantly by interactions between music training and other parameters such as instrumental timbre and register, as shown in Figures 12 and 13, respectively, provides evidence that sound mass perception is dependent both on subjective context and objective properties of the musical stimulus. This complicates the matter of searching for threshold values of sound mass fusion, which may not only covary across interactive parameters, but may also exhibit considerable intersubjective variation.

#### General Discussion and Conclusions

The observations made in these two experiments support our initial suspicion that it is no simple task to say what is or isn't a sound mass, in spite of the matter-of-fact way the term is sometimes used in the musicological literature. Nor is it easy to identify absolute values in sonic parameters that correlate with sound mass perception, because such parameters can covary and

because auditory streaming can shift attention to values in the same parameter at a different structural level. This tendency was demonstrated in our experiments when emergent rhythm, which supplanted overall attack rate as the most perceptually salient feature of temporal organization, was correlated with lower sound mass ratings even when dense and complex pitch structures were present. Many factors are involved in the perceptual fusion that subsumes many sounds into a single mass, including (at least) attack rate, pitch organization, spectral content, register, and the music training of the listener, and in some cases these factors can be interactive. Perhaps this daunting complexity explains in part why authors have preferred to address sound mass at a theoretical or poetic level, avoiding empirical examination.

Many more experiments will be required to provide a comprehensive account of sound mass perception. As noted above, sound mass integration has been achieved in many different ways that cannot be demonstrated in any single piece or compositional method. In some cases, as in *Continuum*, complexity in a parameter that might otherwise lead to sound mass integration (such as pitch complexity) can actually decrease the likelihood of integration occurring due to its interaction with other musical or psychoacoustic parameters (such as emergent rhythm). Nevertheless, based on our experimental observations, some consistent tendencies may be noted. The temporal threshold of sound mass fusion, described by Ligeti (1970) as the point at which tones “merge into a continuum” (p. 1) and by Truax (1988) as “a fusion of grains into a continuous texture” (p. 18), varies with emergent rhythm resulting from pitch organization, and to a lesser extent with register, timbre, and music training, and therefore is not an absolute value of attack rate as has sometimes been suggested.

When attack rate is held constant, as it was in our first experiment, average sound mass ratings varied continuously over the course of the piece, appearing to respond more closely to rhythmic organization than to pitch organization. Indeed, one of the highest-rated sections (H2 above) had one of the broadest compasses and most widely spaced intervallic structure in the piece, challenging the intuitive assumption that higher pitch density is necessarily correlated with higher sound mass fusion. Timbre also appears to be a relevant factor, with ratings for the three instruments (harpsichord, piano, and organ) varying not only in overall value but also in contour. However, these differences are not very great, and further research is necessary before strong conclusions about the nature and extent of the contribution of timbre to sound mass perception can be drawn.

When isolated excerpts with selectively modified parameters were ranked singularly rather than continuously in our second experiment, interactions between parameters became clearer. The already noted variation correlated with different instrumental timbres appears to be stronger when there are cues to promote auditory streaming, such as asynchronous or unequal cyclical patterns, and “outlier” pitches whose registral separation or position at the apex or nadir of relatively longer patterns marks them for segregation. Both attack rate and octave transposition were correlated with a roughly linear pattern, with increasing sound mass ratings observed with increasing speed or decreasing register. These factors seemed to operate independently of one another, but both varied according to the musical organization of the excerpts. Finally, the subjective context of the listener was seen to be an influential factor as participants with little or no music training gave higher ratings overall than musicians, with this tendency being exaggerated for the harpsichord timbre and for lower octave transpositions.

Ligeti himself rarely if ever used the term *sound mass* but described his compositional methods and goals in vivid poetic terms. For example:

All in all, you cannot hear my music as it appears on paper . . . The technical process of composition is like letting a crystal form in a supersaturated solution. The crystal is potentially there in the solution but becomes visible only at the moment of crystallization . . . My aim was to arrest the process, to fix the supersaturated solution just at the moment before crystallization (1983, pp. 14-15)

In *Continuum*, it might be intuitive to assume that “crystallization” corresponds to the emergence of clearly defined intervals (the so-called Ligeti “signals”; Bernard, 1999), whereas more complex or ambiguous structures correspond to supersaturation, which we might equate with mass. This would also be consistent with the descriptions of Ligeti himself, and of several secondary authors (e.g., Bernard, 1999; Clendinning, 1993; Roig-Francolí, 1995). It may thus seem surprising that our data suggest exactly the opposite conclusion: sound mass ratings were highest when the pitch structure was most clear. There are several possible explanations for this. It may be that, in spite of the emphasis on intervals in much of the scholarship, in this instance perceptible rhythm is a more important factor for auditory stream segregation than pitch structure and hence the dissolution of sound mass integration. When rate of attack is sufficiently high and no perceptible rhythm emerges from the musical organization, the conditions

for many sound events to coalesce into a single auditory unit would seem to be satisfied. When the pitch organization changes so that perceptible rhythm does emerge, the sense of perceptual integration is weakened even though the clarity of pitch intervals is subverted (or “blurred”; Bernard, 1999; Roig-Francolí, 1995).

Another possibility is that the definition of sound mass that was given to participants led them to equate total integration with sound mass, and that the simpler pitch structures are more likely to be perceived as totally integrated. This may or may not be problematic. The relation between perceptual integration and sound mass is difficult to pin down. Clearly there must be some positive correlation between the two, for multiple sound events or sources must be grouped perceptually in order for sound mass to have any phenomenal existence. However, this integration can only go so far until it becomes so complete that the perception of multiplicity is nullified and the mass becomes a true singularity, not categorically different from ordinary sound events or auditory streams. It may be argued that the simplest pitch structures in *Continuum* elicited such high responses because they cross that line, becoming too completely integrated to be considered masses. But just where that line might be, in what sonic parameters or values it might consist, whether or not it can be defined in general terms abstracted from particular instances – these questions have yet to be answered, and require much future research.

This research could involve similar experiments on other pieces from the sound mass canon, or on artificially created stimuli, to determine the extent to which the tendencies here noted are generalizable, and to determine the factors at play in sound mass pieces of very different construction. Of particular interest might be pieces with many simultaneous parts, such as Ligeti’s micropolyphonic or Xenakis’s stochastic works, which could be used to evaluate the effect of Huron’s (2001) principles of limited density and part overlapping on sound mass perception. Also of interest would be pieces from later movements influenced by these early sound mass composers, including so-called “spectral” composers such as Gérard Grisey and Tristan Murail who also rely on perceptual fusion but achieve it through more explicitly harmonic and timbral means, composers such

as Kaija Saariaho and Helmut Lachenmann who exploit the continuum between tone and noise in the effecting of perceptual fusion, granular synthesis composers such as Barry Truax and Curtis Roads who invoke sound mass fusion in an acousmatic context, and “saturalist” composers such as Raphael Cendo and Franck Bedrosian for whom perceptual saturation is an explicit goal.

In addition to uncovering whatever commonalities there may be in the perceptual principles, parameters, and threshold values of sound mass fusion in these various genres, it would also be interesting to explore the extent to which they are distinct from one another. This may involve exploration of semiotic as well as perceptual dimensions. As noted above, the everyday concept of *mass* lends itself to various connotative mappings that are not necessarily synonymous, for example formlessness, spatiality, materiality, magnitude, numerousness, and impenetrability. Future research could examine whether or not such connotations are communicated by sound mass music, and whether or not certain approaches to sound mass composition are particularly suited to certain connotations. This would help us begin to understand the extent to which the compelling perceptual and conceptual goals so eloquently described by sound mass composers cross the gulf from poietic to esthetic (Nattiez, 1987/1990), in the hopes of eventually developing a catalogue of sound mass types that would elucidate the artistic significance of this kind of perceptual fusion.

#### Author Note

Chelsea Douglas and Jason Noble contributed equally to the research and writing of this article. This work was supported by a grant from the Fonds de recherche du Québec–Société et culture awarded to SMc, as well as SMc’s Canada Research Chair. We would like to acknowledge the help of Bennett K. Smith in programming PsiExp for Experiment 2.

*Correspondence concerning this article should be addressed to* Chelsea Douglas (chelsea.douglas@mail.mcgill.ca) or Jason Noble (jason.noble@mail.mcgill.ca), Schulich School of Music, McGill University, 555 Sherbrooke St. W., Montreal, Quebec, Canada H3A 1E3.

#### References

- BATES, D., MAECHLER, M., BOLKER, B., & WALKER, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1-48. doi:10.18637/jss.v067.i01.
- BAAYEN, R. H., DAVIDSON, D. J., & BATES, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390-412.



- BERNARD, J. (1999). Ligeti's restoration of interval and its significance for his later works. *Music Theory Spectrum*, 21(1), 1-31.
- BREGMAN, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.
- BROWN, J. (2014). "Audio-visual palimpsests: Resynchronizing silent films with 'special' music." In D. Neumeyer (Ed.), *The Oxford handbook of film music studies* (pp. 588-610). Oxford, UK: Oxford University Press.
- CAMBOUROPOULOS, E., & TSOUGRAS, C. (2009). Auditory streams in Ligeti's *Continuum*: A theoretical and perceptual approach. *Journal of Interdisciplinary Music Studies*, 3(1-2), 119-137.
- CLENDINNING, J. (1993). The pattern-meccanico compositions of György Ligeti. *Perspectives of New Music*, 31(1), 192-234.
- EDWARDS, J. M. (2001). North America since 1920. In K. Pendle (Ed.), *Women and music: A history* (pp. 314-386). Bloomington, IN: Indiana University Press.
- GRISEY, G. (1984). La musique: Le devenir des sons [Music: Sounds becoming]. *Darmstädter Beiträge zur Neuen Musik*, 19, 16-23.
- HURON, D. (2001). Tone and voice: A derivation of the rules of voice-leading from perceptual principles. *Music Perception*, 19, 1-64.
- ISO 389-8. (2004). *Acoustics – Reference zero for the calibration of audiometric equipment – Part 8: Reference equivalent threshold sound pressure levels for pure tones and circumaural earphones* (Tech. Rep.). Geneva, Switzerland: International Organization for Standardization.
- LIGETI, G. (1970). *Continuum*. Mainz: B. Schott's Söhne (Composed in 1968).
- LIGETI, G. (1983). *György Ligeti in Conversation with Péter Várnai, Josef Häusler, Claude Samuel and himself*. London: Eulenburg.
- LIGETI, G. (1997). *Coulée* [Recorded by Z. Szathmáry]. On *György Ligeti: Keyboard Works* [CD]. New York: Sony Classical. (Recorded 1995).
- LONDON, J. (2004). *Hearing in Time: Psychological aspects of musical meter*. Oxford, UK: Oxford University Press.
- MARTIN, F. N., & CHAMPLIN, C. A. (2000). Reconsidering the limits of normal hearing. *Journal of the American Academy of Audiology*, 11(2), 64-66.
- MCADAMS, S. (2014, November). *Toward a psychological foundation for a theory of orchestration* [Lecture]. 2014 William Poland Lecture in Music Theory, Ohio State University.
- NATTIEZ, J.-J. (1990). *Music and discourse: Toward a semiology of music*. (C. Abbate, Trans.). Princeton, NJ: Princeton UP. (Original work published 1987)
- NUSSBAUM, C. (2007). *The musical representation: Meaning, ontology, and emotion*. Cambridge, MA: MIT Press.
- PEITGEN, H.-O. (2011). Continuum, chaos and metronomes – A fractal friendship. In L. Duchesneau & W. Marx (Eds.), *György Ligeti: Of foreign lands and strange sounds* (pp. 87-106). Woodbridge, UK: Boydell Press.
- PETERSEN, P. (2008). *Jede zeitliche Folge von Tönen, Klängen, musikalischen Gestalten hat Rhythmus. Über die Rhythmik in Ligetis Cembalostück Continuum* [Each time sequence of tones, sounds, musical forms has rhythm. On rhythm in Ligeti's harpsichord piece Continuum]. Retrieved from [www.saitenspiel.org/saitenspiel.org/Ligeti\\_Continuum.html](http://www.saitenspiel.org/saitenspiel.org/Ligeti_Continuum.html) November 30, 2014.
- REDLINGER, E. (2008). *Mrmr* [Software]. Retrieved December 2, 2014 from <http://opensoundcontrol.org/implementation/mrmr>
- ROADS, C. (2003). The perception of microsound and its musical implications. *Annals of the New York Academy of Sciences*, 999(1), 272-281.
- ROIG-FRANCOLI, M. (1995). Harmonic and formal processes in Ligeti's net-structure compositions. *Music Theory Spectrum*, 17(2), 242-267.
- SMALLEY, D. (1993). Defining transformations. *Interface*, 22(4), 279-300.
- SMITH, B. K. (1995). PsiExp: An environment for psychoacoustic experimentation using the IRCAM musical workstation. In D. Wessel (Ed.), *Proceedings of the Society for Music Perception and Cognition Conference'95*. Berkeley, CA: University of California, Berkeley.
- TREMBLAY, A., & RANSIJN, J. (2013). A suite of functions to back-fit fixed effects and forward-fit random effects, as well as other miscellaneous functions. R package version 2.5. <https://www.researchgate.net/publication/235709549> (last accessed Dec 9, 2015) [Computer software]
- TRUAX, B. (1988). Real-time granular synthesis with a digital signal processor. *Computer Music Journal*, 12(2), 14-26.
- WARREN, R. (2008). *Auditory perception: An analysis and synthesis* (3rd ed.). New York: Cambridge University Press.
- WEST, B. T., WELCH, K. B., & GALECKI, A. T. (2007). *Linear mixed models: A practical guide using statistical software*. Boca Raton, FL: Chapman & Hall/CRC Press.
- XENAKIS, I. (1971). *Formalized music: Thought and mathematics in composition*. Bloomington, IN: Indiana University Press.
- ZBIKOWSKI, L. (2002). *Conceptualizing music: Cognitive structure, theory, and analysis*. Oxford, UK: Oxford University Press.