

# Journal of Experimental Psychology: Human Perception and Performance

## Linking Melodic Expectation to Expressive Performance Timing and Perceived Musical Tension

Bruno Gingras, Marcus T. Pearce, Meghan Goodchild, Roger T. Dean, Geraint Wiggins, and Stephen McAdams

Online First Publication, November 23, 2015. <http://dx.doi.org/10.1037/xhp0000141>

### CITATION

Gingras, B., Pearce, M. T., Goodchild, M., Dean, R. T., Wiggins, G., & McAdams, S. (2015, November 23). Linking Melodic Expectation to Expressive Performance Timing and Perceived Musical Tension. *Journal of Experimental Psychology: Human Perception and Performance*. Advance online publication. <http://dx.doi.org/10.1037/xhp0000141>

# Linking Melodic Expectation to Expressive Performance Timing and Perceived Musical Tension

Bruno Gingras  
University of Innsbruck

Marcus T. Pearce  
Queen Mary University of London

Meghan Goodchild  
McGill University

Roger T. Dean  
University of Western Sydney

Geraint Wiggins  
Queen Mary University of London

Stephen McAdams  
McGill University

This research explored the relations between the predictability of musical structure, expressive timing in performance, and listeners' perceived musical tension. Studies analyzing the influence of expressive timing on listeners' affective responses have been constrained by the fact that, in most pieces, the notated durations limit performers' interpretive freedom. To circumvent this issue, we focused on the unmeasured prelude, a semi-improvisatory genre without notated durations. In Experiment 1, 12 professional harpsichordists recorded an unmeasured prelude on a harpsichord equipped with a MIDI console. Melodic expectation was assessed using a probabilistic model (IDyOM [Information Dynamics of Music]) whose expectations have been previously shown to match closely those of human listeners. Performance timing information was extracted from the MIDI data using a score–performance matching algorithm. Time-series analyses showed that, in a piece with unspecified note durations, the predictability of melodic structure measurably influenced tempo fluctuations in performance. In Experiment 2, another 10 harpsichordists, 20 nonharpsichordist musicians, and 20 nonmusicians listened to the recordings from Experiment 1 and rated the perceived tension continuously. Granger causality analyses were conducted to investigate predictive relations among melodic expectation, expressive timing, and perceived tension. Although melodic expectation, as modeled by IDyOM, modestly predicted perceived tension for all participant groups, neither of its components, information content or entropy, was Granger causal. In contrast, expressive timing was a strong predictor and was Granger causal. However, because melodic expectation was also predictive of expressive timing, our results outline a complete chain of influence from predictability of melodic structure via expressive performance timing to perceived musical tension.

*Keywords:* performance, tension, expectations, entropy, information content, timing

Music is a uniquely human and cross-culturally universal means of conveying information using sequential streams of auditory events (Clarke, 1985, 1988; Palmer, 1989, 1996). Although there are major differences between language and music as modes of communication, both can be viewed as a means of transferring information (Koelsch et al., 2004). As in linguistic communica-

tion, musical communication involves the elaboration of a musical message (composition), its utterance (performance), and finally its perception and cognitive appraisal by one or more recipients (listeners). Music is an important part of many people's everyday lives (Sloboda, 2010), and the information it conveys can profoundly affect the mood, emotions, and well-being of listeners

---

Bruno Gingras, Institute of Psychology, University of Innsbruck; Marcus T. Pearce, School of Electronic Engineering & Computer Science, Queen Mary University of London; Meghan Goodchild, Schulich School of Music, McGill University; Roger T. Dean, MARCS Institute, University of Western Sydney; Geraint Wiggins, School of Electronic Engineering & Computer Science, Queen Mary University of London; Stephen McAdams, Schulich School of Music, McGill University.

Supported by a Social Sciences and Humanities Research Council postdoctoral fellowship (Award 756-2008-0289; Bruno Gingras), the Canadian Natural Sciences and Engineering Research Council (Grant RGPIN 312774-2010; Stephen McAdams), a Canada Research Chair (Stephen McAdams), and the

Engineering and Physical Sciences Research Council (Grant EP/H01294X/1; Marcus T. Pearce and Geraint Wiggins). We thank Alain Poirier, former director of the Conservatoire National Supérieur de Musique et de Danse de Paris, for the loan of the MIDI harpsichord, and Lauren Stewart, for use of the sound-isolation chamber at Goldsmiths (London). We are grateful to Bennett K. Smith for programming and helping design the experimental interface. We also thank Pierre-Yves Asselin, Julien Boissinot, and Harold Kilianski for their technical assistance; McGill Sound Recording for permission to use their equipment; and Martha de Francisco for her assistance with sound recording.

Correspondence concerning this article should be addressed to Bruno Gingras, Institute of Psychology, University of Innsbruck, Innrain 52f, Innsbruck, Austria A-6020. E-mail: brunogingras@gmail.com

(Panksepp & Bernatzky, 2002). As such, research on music cognition broadens our perspective on human cognition in general by considering aspects that share functionalities with other types of cognition, but more importantly by considering its specificities.

Kendall and Carterette's (1990) theoretical model of musical communication, based on traditional Western art music, expresses this three-way relationship between the composer, the performer(s), and an audience. Following this model, we posited that a performer's interpretation of a musical composition (the message) is influenced by the structure of the composition, and that the listener's response to a particular performance of a piece reflects both the musical structure as notated in the score and the specific features of the performer's interpretation.

The purpose of this article was to study the processes involved in the musical flow of information from the composer through the performer to the listener. The vast majority of empirical research in the field has focused exclusively on one or two of these aspects of musical communication (e.g., Gabrielsson & Juslin, 1996; Juslin & Timmers, 2011). Here we explicitly attempted to study empirically all three components in the communicative chain. In doing so, we first asked *what* is communicated and second, *how* it is communicated. Our primary aim in this study was to address the latter question. First, however, we must establish our assumptions about the content of musical communication. In doing so, we begin with the listener and then consider the structure of the composition and finally the role of the performer.

### The Listener: Emotion, Tension, and Expectation

Previous research has suggested that listeners value music primarily for its impact on mood and emotion (Juslin & Laukka, 2004) and the content of musical communication is thought to be primarily affective rather than referential (Hanslick & Cohen, 1854; Meyer, 1956). There exist several mechanisms through which music can have an effect on the emotional state of listeners (Juslin & Västfjäll, 2008). Of these, we will focus on expectation as a vehicle for studying the communication of emotion by music because it is more directly based on the structure of the music than other mechanisms (e.g., episodic memory or evaluative conditioning; see Juslin & Västfjäll, 2008), thereby facilitating the establishment of a link between compositional structure, performance, and perception. Musical structures generate expectations for how the music will continue, and these expectations change dynamically as the musical structure unfolds in time. At any given moment, the music may confirm or violate the expectations of a given listener to different degrees. Real-time responses to music can be used to measure this continuous process (Krumhansl, 1996, 2002; Toiviainen & Krumhansl, 2003; McAdams, Vines, Vieillard, Smith, & Reynolds, 2004). In particular, the perception of musical tension, including its connection to aesthetic responses by listeners, has been studied extensively (Farbood, 2012; Fredrickson, 1995, 2000; Madsen, 1997; Madsen, Brittin, & Capperella-Sheldon, 1993; Madsen & Fredrickson, 1993).

Expectation is thought to be an important cognitive mechanism in language processing (Cristià, McGuire, Seidl, & Francis, 2011; DeLong, Urbach, & Kutas, 2005; Hale, 2006; Levy, 2008; Saffran, 2003), visual perception (Bar, 2007; Bubic, von Cramon, & Schubotz, 2010; Egner, Monti, & Summerfield, 2010), and motor sequencing (Wolpert & Flanagan, 2001). It is also an important

psychological process determining the aesthetic experience of music listeners (Hanslick & Cohen, 1854; Huron, 2006; Meyer, 1956). In particular, violations of expectation are thought to lead to increases in tension (Meyer, 1956), as reflected in increased physiological arousal (Egermann, Pearce, Wiggins, & McAdams, 2013; Steinbeis, Koelsch, & Sloboda, 2006). Analysis of behavioral ratings of tension experienced while listening to Mozart's Piano Sonata in E-flat major, K. 282 (Krumhansl, 1996, 2002) showed that tension ratings covaried both with surface features, such as pitch height (register), and features associated with the stability of a pitch within a given tonal framework (tonal stability) or the tendency for certain pitches in a melody to move toward other pitches to a greater or lesser extent (tonal attraction) (Lerdahl, 2001). The latter features are likely to contribute to an experience of expectation violation in listeners enculturated with Western tonal music (Lerdahl & Krumhansl, 2007). However, a full model of musical expectation depends not only on the degree of tonal stability, but also on the schematic effects of long-term musical exposure and on the presence of repeated or elaborated structures (e.g., motifs, theme, variation) within a musical work (Pearce, 2005). The goal of the present research was to examine the hypothesized relation between the expectations generated by such a model and behavioral ratings of tension in musical listening.

### The Composition: Musical Structure, Predictability, and Expectation

Expectation can be viewed as a process of probabilistic prediction based on knowledge of musical structure acquired through statistical learning (Huron, 2006; Meyer, 1956; Pearce, 2005; Pearce & Wiggins, 2012). Note that similar processes are thought to operate in language perception (Saffran, 2003; Cristià et al., 2011; Hale, 2006; Levy, 2008) and other areas of perceptual processing (Friston, 2010; Perruchet & Pacton, 2006). There is empirical evidence that listeners acquire expectations for musical events through implicit statistical learning of regularities through long-term exposure to music (Krumhansl, 1990), and listeners' expectations also reflect the learning of statistical properties of the local context (Oram & Cuddy, 1995; Saffran, Johnson, Aslin, & Newport, 1999; Tillmann, Bharucha, & Bigand, 2000; Tillmann, Bigand, & Pineau, 1998). This process of statistical learning allows listeners to generate probabilistic predictions about forthcoming musical events, dependent on the prior musical context and previously acquired schematic expectations for the musical style in question (Krumhansl, Louhivuori, Toiviainen, Järvinen, & Eerola, 1999; Pearce, Ruiz, Kapasi, Wiggins, & Bhattacharya, 2010; Pearce & Wiggins, 2006). Consistent with an approach based on statistical learning, melodic pitch expectations vary between musical styles (Krumhansl et al., 2000) and cultures (Carlsen, 1981; Castellano, Bharucha, & Krumhansl, 1984; Eerola, 2004; Kessler, Hansen, & Shepard, 1984; Krumhansl et al., 1999), throughout development (Schellenberg, Adachi, Purdy, & McKinnon, 2002), and across degrees of musical training and familiarity (Krumhansl et al., 2000; Pearce et al., 2010).

In this study, we used a computational model of auditory expectation, IDyOM (Information Dynamics of Music), to specify precise, quantitative measures of structural predictability for each note in a melody. The model itself has been presented (Conklin & Witten, 1995; Pearce, 2005) and tested (Egermann et al., 2013;

Hansen & Pearce, 2014; Omigie, Pearce, & Stewart, 2012; Omigie, Pearce, Williamson, & Stewart, 2013; Pearce, Conklin, & Wiggins, 2005; Pearce et al., 2010; Pearce & Wiggins, 2004, 2006), so we restrict ourselves to a summary of features most relevant for the present study.

The central feature of the model is that it learns about sequential dependencies between notes in an unsupervised manner through exposure to melodies. In the context of this research, the model focuses specifically on the pitch of each note in a melody. Given such a representation of music, the model returns a conditional probability distribution of pitches for each note, given the preceding sequence of pitches. Two information-theoretic measures are derived from the predictive distribution generated for each note: *entropy* is a measure of the uncertainty of the predictive distribution associated with the melodic context (maximum entropy occurs for a distribution in which every pitch is equally likely to follow the context) whereas *information content* (IC; the negative log probability) is a measure of the unexpectedness of the note that actually follows.

Entropy reflects the uncertainty of the prediction before the next note is heard, whereas IC reflects how unexpected the next note is once it actually happens. If entropy is high, then every possible continuation of the sequence is equally likely to the model (i.e., it is very uncertain about the continuation). Alternatively, a very low entropy signifies that one note is highly expected, and other possibilities are very unexpected (i.e., the model is very certain). In this latter situation, if the certain prediction is correct, then the next note is highly expected and the IC is low; if, however, the certain prediction is incorrect, then the next note is unexpected and the IC is high. Therefore, accurate predictive models are associated with both low entropy predictions and low IC outcomes. For example, a listener who is sure that the pitch of the next note in a melody will be a middle C (261.63 Hz) and not any other note has low predictive uncertainty at that point in time (before the next note is heard); another listener to the same melody who expects to hear any pitch between middle C and the A above it (440 Hz) with equal likelihood has higher predictive uncertainty. Note that predictive uncertainty is distinct from the expectedness of the event that actually happens. Our first listener is highly certain about the next event in the melody, but may be incorrect (and surprised) if the note that actually follows is not a middle C.

The model is based on a Markov or  $n$ -gram model (Manning & Schütze, 1999, chap. 9), which computes the conditional probability of a note given the  $n - 1$  preceding notes in the melody. The quantity  $n - 1$  is called the order of the model, and basic Markov models have a fixed order (as a result, they suffer from the problem that previously unseen notes are estimated with a probability of zero). Basic Markov modeling is extended in two ways. First, the model is of variable order, computing a weighted average of probabilities from models of different order. This allows the system to benefit both from the structural specificity of longer (but relatively rare) contexts and the statistical power of more frequent (but less specific) low-order contexts. It also allows the model to generate probabilities for notes that have yet to appear in a given context.

Second, because perceptual representations of music are multi-dimensional (e.g., Levitin & Tirovolas, 2009; Shepard, 1982), the system has the ability to use a combination of different features, or *viewpoints*, to predict the properties of notes (Conklin & Witten,

1995; Pearce et al., 2005). For example, in predicting the pitch of the next note, the system may combine models trained on representations of *chromatic pitch*, *scale degree* (the position of the pitch in the tonal scale), *pitch interval* (the distance in pitch between two consecutive notes), or *contour* (whether the pitch rises, falls, or remains the same). Two viewpoints may be linked ( $A \times B$ , where A and B are the source viewpoints), so that notes are represented by pairs of values. For example, a link between pitch interval and scale degree would represent each note as a pair of values representing, respectively, the pitch interval and the scale degree of that note.

A separate model is used for each viewpoint specified, each of which generates a probability distribution. These distributions are subsequently combined using a weighted geometric mean into a single probability distribution governing the pitch of the next note. The viewpoints used may be specified by hand or optimized using a hill-climbing search through the space of possible sets of viewpoints for sets that minimize the IC (or maximize the predictability) of the music being studied. The hill-climbing procedure starts with an empty set of viewpoints and, at each step, makes the single addition or deletion of a viewpoint, if any exist, which reduces the average IC of the musical piece by the greatest amount (Pearce, 2005).<sup>1</sup>

In this research we used two sets of viewpoints: first, a single *default* linked viewpoint combining pitch interval and scale degree (i.e., each note is represented as a pair of values—the first reflects the pitch interval in semitones from the previous note, the second reflects the chromatic scale degree from 0 [*the tonic*] to 11 [*the leading tone*]). Second, we examined an *optimized* model in which viewpoints were selected through hill-climbing. The entire set of viewpoints considered and the optimized viewpoints selected are listed in the Appendix.

Regardless of the model's configuration, its output is always a conditional probability distribution governing the pitch of the next note in a melody. As noted above, the entropy of this distribution reflects the predictive uncertainty of the model, whereas the IC (the negative log probability) of the note that actually follows reflects the unexpectedness of that note.

### The Performer: Improvisation and Expressive Performance Timing

We turn now to the performer, the expressive intermediary between the musical structures created by the composer and the acoustic signal perceived by the audience. A substantial body of research on music performance has shown that performers' interpretive choices are informed by the formal structure of the musical pieces. For instance, it is well-established that performers, like speakers (Cooper & Danly, 1981), tend to slow down at sectional boundaries, a process called "phrase-final lengthening" (Clarke, 1989; Palmer, 1989; Repp, 1990; Todd, 1985), and that the magnitude of this deceleration tends to correspond to the hierarchical importance of the boundaries (Shaffer & Todd, 1987). Similarly,

<sup>1</sup> It is possible that, during the hill-climbing procedure, the addition of subsequent viewpoints renders an existing one redundant or incompatible, so deleting it increases prediction performance (reducing information content). This was not the case for the optimized model we used herein (but see Pearce, 2005, p. 163, for an example).

the music communication hypothesis (Clarke, 1985, 1988; Palmer, 1989, 1996; Repp, 1992, 1995) suggests that tempo fluctuations are a means of conveying information about the grouping structure of a piece. Thus, the performer plays a crucial role in clarifying the musical structure for the audience.

Although tempo fluctuations on both local and large-scale levels have been identified as a major expressive device in music performance, most pieces in the Western repertoire specify a regular rhythmic structure, which greatly restricts the performers' latitude regarding note durations. Consequently, the link between information-theoretic measures derived from the music score and performance timing cannot be analyzed optimally. A few studies have addressed this limitation by focusing on improvised performance (e.g., Large, Palmer, & Pollack, 1995), but this has the drawback, from the viewpoint of Kendall and Carterette's (1990) model of musical communication, of amalgamating the composer and performer functions, making it impossible to isolate the individual characteristics associated with each component. To circumvent this issue and to investigate performance-induced effects on listeners' aesthetic responses, we focused on one particular genre of lute and harpsichord compositions, cultivated especially in France between 1650 and 1720, known as the *unmeasured prelude*. The unmeasured prelude is a semi-improvisatory genre that does not specify a rigid metrical framework—a pattern of strong and weak beats—and even leaves the durations of individual notes unspecified in many instances. It thus allows for a much wider variability in terms of expressive timing (i.e., variations in tempo made by a performer, which may differ between performances) than most Western musical genres (Gingras, Asselin, & McAdams, 2013), providing an ideal repertoire for investigating the links between information-theoretic measures derived from the musical score, expressive timing in performance, and listeners' aesthetic responses.

Following the musical communication hypothesis described above, we hypothesized that performers (like listeners) are influenced by the predictability of the music they are performing. In perception, it has been shown that unexpected events take longer to process than expected ones (Bharucha & Stoeckig, 1986, 1987; Omigie et al., 2012; Pearce et al., 2010). We predicted that this effect would extend to performance such that structural expectations influence expressive timing. In particular, we hypothesized that unexpected structures would be associated with a slowing of tempo and with a greater variability in expressive timing. Our

decision to study harpsichord performance was also motivated by a desire to minimize the role of acoustic intensity changes, because they are thought to be major influences on perception of affect and tension (Ilie & Thompson, 2006; Krumhansl, 1996). Because the harpsichord has a relatively small dynamic range (Fletcher, 1977; Penttinen, 2006), particularly with respect to the performance of a particular note or chord, we expected that the nature and roles of performance timing would be readily detectable.

## The Present Study

The purpose of the present research was to elucidate the cognitive mechanisms engaged in the model of musical communication proposed by Kendall and Carterette, involving the flow of information from the composer via the performer to the listener. More precisely, the cognitive mechanism we sought to elucidate is *expectation*, which we conceived as a dynamic process of predicting future elements in a sensory stimulus using both the preceding context within the stimulus and the prior experience of the observer (Bubic et al., 2010). We simulated expectation formally in terms of the information-theoretic quantities IC (reflecting the unexpectedness of an event) and entropy (the predictive uncertainty about which event will follow, before it appears). We derived these quantities using a computational model of auditory expectation (IDyOM), which has been shown to account accurately for listeners' pitch expectations in music (Egermann et al., 2013; Hansen & Pearce, 2014; Pearce et al., 2010; Pearce & Wiggins, 2006; Omigie et al., 2012, 2013). In this study, we examined the role of expectation, as a cognitive process, in musical communication. Specifically, we hypothesized, on the basis of previous research (Meyer, 1956; Narmour, 1990; Huron, 2006), that composers would introduce patterns that confirm or disconfirm expectations to different degrees, that performers would perceive this ebb and flow of predictability and reflect it in their expressive performance, and that listeners, in turn, would experience heightened tension and relaxation in expressive performances at points where the process of expectation is, respectively, disrupted (by high unexpectedness or uncertainty) or stabilized (by low unexpectedness or uncertainty).

We conducted our investigation in two studies (Experiments 1 and 2) for which we report time-series analyses (TSAs) of continuous processing of an unmeasured prelude for harpsichord by Louis Couperin (1626–1661), the *Prélude non mesuré No. 7* (a

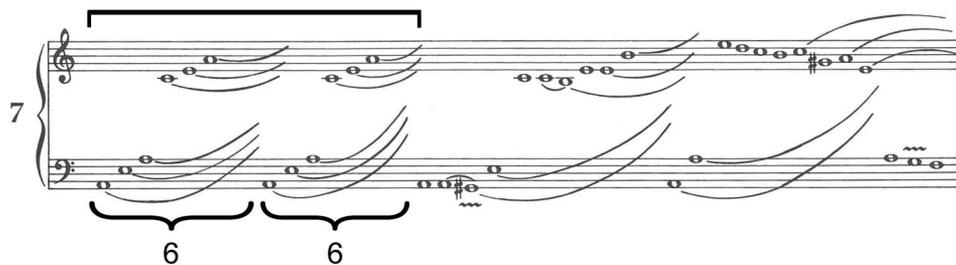


Figure 1. Excerpt from the score of the *Prélude non mesuré No. 7*, by Louis Couperin (1626–1661). The excerpt corresponds to the first of five musical systems. The opening arpeggio pattern (with its 6 + 6 subdivision) is indicated by brackets. Adapted from *The art of the French unmeasured prelude*, C. Tilney (Ed.), 1991, Vol. 2. Copyright 1991 by Schott.

score excerpt is provided in Figure 1). This prelude comprises a single unaccompanied melodic line in which pitches unfold successively, although harmonies are sometimes implied through the use of arpeggiation (a melodic succession creating the semblance of chords) and sustained notes.<sup>2</sup>

Experiment 1 examined relations between the changing structure of the music (specifically, structural predictability as assessed by the computational model of auditory expectation described above) and expressive timing in performances of the piece. Experiment 2 examined the impact of both structural predictability and expressive timing on the emotional experience of listeners, as measured by continuous tension ratings. As mentioned, the unmeasured prelude is an ideal genre to address these issues, due to the absence of a rigid metrical framework.

Melodic expectations, tempo variations in performance, and perceived musical tension exemplify the three vertices of a relation involving musical structure, expressive strategies, and aesthetic responses. The present study focused on these three parameters, aiming to quantify the individual components and to assess the nature and directionality of their relations with each other. We hypothesized that unexpected events (those with high IC) and points of predictive uncertainty (high entropy) would be associated with a slowing down of performance timing and greater variability between performers. In turn, we hypothesized that high IC and high entropy, in concert with their influence on performance timing, would be associated with a greater perceived tension in listeners.

## Experiment 1: Performance Timing

### Method

**Participants and procedure.** Twelve professional harpsichordists (five female and seven male) participated in the study. They were professional harpsichordists from the Montreal (Canada) area or harpsichord students at McGill University in Montreal. Their average age was 39 years (range: 21–61 years). They had played the harpsichord for a mean duration of 22 years (range: 6–40 years). Seven had previously won prizes in regional, national, or international harpsichord competitions.

The harpsichordists were provided with the score of the *Prélude non mesuré No. 7* by Louis Couperin 4–6 weeks prior to the recording session. They performed the piece twice (with no special directives) using the score. Performances took place in an acoustically treated studio on an Italian-style Bigaud harpsichord (Paris, Heugel) with two 8-foot stops. Only the back stop was used for the experiment. This harpsichord was equipped with a MIDI console, allowing precise measurement of performance parameters (note onset/offset and key press velocity).

The audio signal was recorded through two omnidirectional microphones (MKH 8020; Sennheiser Electronic GmbH, Wedemark, Germany). The microphones were located 1 m above the resonance board and were placed 25 cm apart. The audio and MIDI signals were sent to a personal computer using an RME Fireface audio interface (Audio AG, Haimhausen, Germany). Audio and MIDI data were then recorded using Cakewalk's SONAR software (Cakewalk, Boston, MA) and stored on a hard disk.

**Analysis of the performance data.** Note onsets and offsets were extracted from the MIDI data of the 24 performances and

matched to the score using an algorithm developed by Gingras and McAdams (2011), which has been shown to be suitable for ornamented harpsichord pieces (Gingras et al., 2013). The mean error rate per performance, defined as the proportion of wrong notes or missing notes relative to the total number of score notes, was 0.54% (range: 0–2.14%). This low error rate is comparable to the rates reported by Repp (1996) and Goebel (2001) in studies on professional piano performance, suggesting that the performance data collected for the current study were of suitable quality for assessing performance timing in professional harpsichord performance.

Tempo patterns were determined using interonset intervals (IOI). IOIs were measured as the time interval between the onsets of successive notes, as obtained from the MIDI data for each performance of the piece. Each of the 140 note events in the *Prélude* was used to generate a series of IOIs. The IOI for the last note of the piece was taken to be the onset–offset time interval. Although this interval was, on average, more than four times as long as the IOI of any other event in the piece, it was found by several additional analyses that excluding the last note from the data did not change optimized models. There were very few note events for which no timing data were available (0.60% of the total number of events): 18 wrong or missing notes, plus 2 complex ornamentations for which a precise onset could not be computed. IOIs for these notes were interpolated by using the average of the data for all the other performers, each note onset in the MIDI data having been previously unambiguously matched to a score event in the piece.

The IOI time series thus obtained for both performances by each harpsichordist yielded 12 within-performer average IOI series that were used in all subsequent analyses in Experiment 1. These within-performer IOI series were then averaged across performers for some analyses. Interperformer variability was assessed event by event to create a time series of the coefficient of variation (CV) of the IOIs (the ratio of the standard deviation to the mean for each event considering all 12 within-performer IOI series).

**Information-dynamic analysis.** The computational model described in the section “The Composition: Musical Structure, Predictability, and Expectation” (Pearce, 2005) was used to generate an information-dynamic analysis of the *Prélude*, including measures of the IC (unexpectedness) and entropy (uncertainty) for each note in context, given the model. We used two separate models with distinct representations. The first used a *default* feature consisting of a single linked viewpoint combining pitch interval and scale degree (i.e., each note was represented as a pair of values—the first reflecting the pitch interval in semitones from the previous note, the second reflecting the chromatic scale degree from 0 [*the tonic*] to 11 [*the leading tone*]). Because the piece is nonmodulatory, the mapping between chromatic pitch and scale degrees was treated as consistent throughout the piece. This viewpoint allowed the model to learn about dependent melodic and tonal relations in the sequential structure of the music. Previous research has found support for the influence of these two factors on pitch expectations in

<sup>2</sup> In terms of musicologically conceived tension, the structure of the piece progresses from maximum stability to maximum instability and back based on its gestural structure (stable arpeggios enclosing less stable melodic passages) (see Goodchild, Gingras, & McAdams, in press).

tonal melody (Hansen & Pearce, 2014; Pearce et al., 2010; Pearce & Wiggins, 2006; Schellenberg, 1997). The second approach used an *optimized* set of viewpoints selected using the hill-climbing procedure described in the section “The Composition: Musical Structure, Predictability, and Expectation” (see also Appendix).

**Time-series analysis.** A detailed presentation of TSA techniques in the study of continuous musical events is available elsewhere (Bailes & Dean, 2012; Dean & Bailes, 2010). This large and long-standing suite of techniques is appropriate for use with data that are serially correlated (thus not independent measures), unlike most common statistical approaches that require independence of the multiple data measures to be studied. TSA is thus a specialized form of regression analysis that can take proper account of autocorrelations, which standard techniques fail to do. Autoregressive (AR) TSA can model the influence of external predictor variables (X), such as the information-dynamic parameters just mentioned, or acoustic properties such as intensity. Data to be modeled need to be statistically stationary, lacking trends. In essence, stationarity means that as the data series fluctuates, its mean and variance remain constant, and the autocorrelation between measures made at any fixed time separation remains constant (as revealed in the so-called autocorrelation function). This is most commonly achieved by differencing the time series, that is, constructing a new series comprised of the differences between successive members of the original (which is thereby one value shorter). For a series called *seriesname*, we refer to this differenced series as *dseriesname*, and of course it can be used to reconstruct the original series, so that models of *dseriesname* also provide predictions for *seriesname* and thus are informative about its predictors. Model selection was based on parsimony and information criteria. Here, we used the Bayesian information criterion (BIC; Schwarz, 1978). A lower BIC value indicates a better model for a given data series (BIC values cannot be meaningfully compared between models of different data series). The absolute value of the BIC is not informative. The quality of a model is assessed by a range of criteria, including the nature of the residual series and the degree of fit of the model to the data. Residuals, corresponding to the variance associated with each data point that is not explained by the model, should themselves lack autocorrelation and be normally distributed around 0 (this is also known as a Gaussian white noise distribution). Otherwise, they still contain significant information that should be modeled. The degree of fit is assessed using the log-likelihood of the model.

We refer to our standard TSAs as ARMAX: autoregressive moving average models with external predictors. We use moving average (MA) components in some cases, but they can be replaced by AR terms if necessary by a reparameterization, and so we do not discuss them in detail. TSA can also assess factors that may have mutual influence, such as perception of loudness and pitch, and consider their possible influences on each other at the same time as their influences on another measured variable. This is done by vector autoregression (VAR), in which multiple variables may be constrained either as exogenous (similar to a psychological independent variable) or endogenous (similar to a dependent variable). Note that all variables were initially included in the VAR analyses; the order of addition and removal of variables has no influence on the final models. This permits a direct comparative assessment of the predictive utility of the selected variables. Additional specific details are provided below as appropriate, and the

reader is referred to standard textbooks on TSA, as well as the summary guidance in Dean and Bailes (2010), Bailes and Dean (2012), and Dean, Bailes, and Dunsmuir (2014a, 2014b). All TSA presented here were conducted using the commercial software STATA (StataCorp, College Station, TX). All statistical analyses involved principled manual processes and decisions; there was no automation of model selection.

## Results

Our goal in Experiment 1 was to examine a hypothesized relation between variations in expressive timing in performance and changes in melodic predictability based on the IC and entropy estimated by the IDyOM model. First, we examine the patterns of IC and entropy; second, we describe how we measured variations in expressive timing; and finally, we use TSA to examine the extent to which changes in IC and entropy account for changes in performance timing.

**Information-theoretic structure of the piece.** We considered the output of two IDyOM models, using *default* and *optimized* viewpoints. The IC and entropy values estimated by the models are shown in Figure 2. It can be seen that the IC is, on average, lower for the default than the optimized viewpoint (although the two are highly correlated). This was true both in absolute terms and in relation to the entropy, indicating that the optimized set of viewpoints allows the model to predict the pitch of the individual notes in the *Prélude* more accurately than the default viewpoint.

**Overview of the performance timing data.** The average duration of a performance (from first to last onset) was 59.9 s for the *Prélude* (range: 39.1–84.2 s). These durations corresponded to a mean tempo of 149.2 events/min (range: 99.8–214.7 events/min). Figure 3 shows the mean IOI for each note event, with the shaded area corresponding to the standard deviation. In contrast to previous studies on measured organ music, which reported high correlations for timing patterns within and between performers (Gingras, 2008), we found a moderate mean correlation of 0.54 ( $SD = 0.18$ ,  $df = 138$ ) between pairs of performances of the *Prélude* played by different performers. Note that even though these correlations may have been influenced by autocorrelation (the tempo was likely to be more similar for nearby note events than for temporally distant ones), the moderate values observed for the *Prélude* imply a fair amount of dissimilarity among the performances. Interestingly, the mean correlation between performances by the same performer reached 0.87 ( $SD = 0.10$ ,  $df = 138$ ). Overall, these analyses suggested that, although both recordings from the same performer generally displayed similar timing patterns, there was a fair amount of variation among performers (see Gingras et al., 2013, for a detailed investigation of the individual variation between performers).

**Prediction of performance timing and interperformer timing variability.** The IOI time series was not statistically stationary (which is required for reliable TSA). However, its first difference (*dIOI*) was stationary, provided that the IOI of the final note of the piece was excluded; despite the exceptional nature of the timing of the final note, its inclusion or exclusion in the time-series models did not change the specifications of the optimal form. Thus, initial modeling of the IOI used the first differenced versions of the relevant data streams, the *dIOI* and those from the IC/

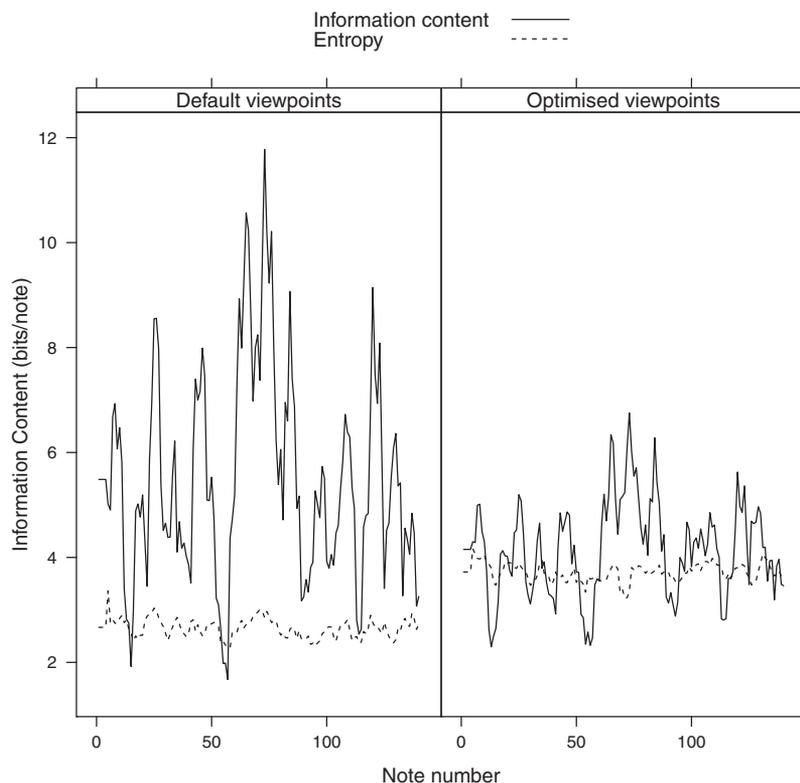


Figure 2. Information content and entropy predicted by the default and optimized versions of the IDyOM (Information Dynamics of Music) model.

entropy analysis in IDyOM. IOI itself showed notable partial autocorrelations at Lags 1, 6, and 7 (corresponding to an autocorrelation between the current note duration and those of notes located 1, 6, and 7 events earlier, respectively). Further analyses showed that the periodicities at 6 and 7 events were related to the

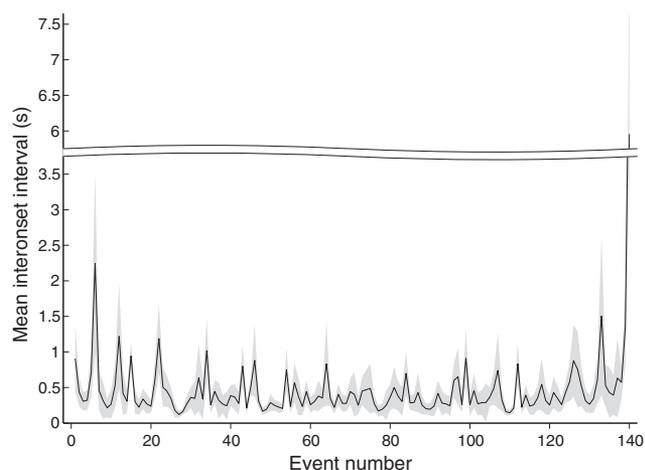


Figure 3. Mean duration (in seconds) of the interonset interval for each event of the *Prélude non mesuré No. 7*. The shaded area corresponds to 1 SD above and below the mean for each event. Note the discontinuity in the vertical axis.

arpeggio figurations found at the beginning and end of the piece. The opening arpeggio pattern contains a 6 + 6 subdivision (see Figure 1), whereas the final arpeggio pattern is characterized by a 6 + 7 subdivision, and these subdivisions were generally clearly delineated by the performers. Moreover, the significant partial autocorrelations at Lags 6 and 7 disappeared when these arpeggio patterns were excluded from the analysis. Graphs of the IOI and IC series showed that long IOIs often anticipate peaks in IC by about one event, whereas they seemed to coincide with troughs in entropy, consistent with our prediction that (statistically) unexpected events, even though they were already known to performers, would be associated with slowing.

Modeling of dIOI indicated that leads (F in Table 1), corresponding to subsequent events of the IDyOM data streams, were more influential than lags (earlier events). For both default and optimized models, dEntropy (the first difference of the entropy time series) was not influential, but Lead 1 of dIC (the first difference of the IC time series) was. Entropy components were considered as predictors in all cases, but were dismissed during the model optimization process on the basis of BIC. In other words, whereas adding a predictor to a model nearly always enhanced its fit, this improvement may not be sufficient to justify the inclusion of an additional predictor to the model (as judged here by the BIC), and hence it was excluded.

Once a good model of a time series is obtained, it shares the autocorrelation features of the data series. Thus measurement of correlation between model and data now becomes informative,

Table 1

*Time-Series Analysis of Performance Timing (dIOI and CV of IOI) Using Information Content and Entropy (From Two IDyOM Models, Using Default and Optimized Viewpoints) as Predictors*

Modeled parameter	Predictors considered (IDyOM model in parentheses)	Components of the optimal model (coefficients for the external predictors in parentheses)	BIC	Correlation coefficients (modeled vs. observed values)
dIOI	IC and entropy (default)	F1.dIC (+.01), AR(1–5)	70.1	.65
dIOI	IC and entropy (optimized)	F1.dIC (+.02), AR(1–5)	71.8	.64
CV of IOI	IC and entropy (optimized)	F1.IC (+.02), F1–2.Entropy (+.06, +.05), AR(1–2), MA(25)	–150.6	.52

*Note.* The chosen data series were stationary, all models listed showed satisfactory white noise residuals. For example, F1.dIC (+.02) (row 4) means: lead one of the (first) differenced IC data series (derived from the optimized IDyOM model in this case) had a coefficient of +.02. See the Method section in Experiment 1 for further details of the time-series analysis. AR(n) = autoregressive error terms; BIC = Bayesian information criterion; CV = coefficient of variation; Fn = lead n; IC = information content; IDyOM = Information Dynamics of Music; IOI = interonset interval; Ln = lag n; MA(n) = moving average error term lag n, where n represents the number of events.

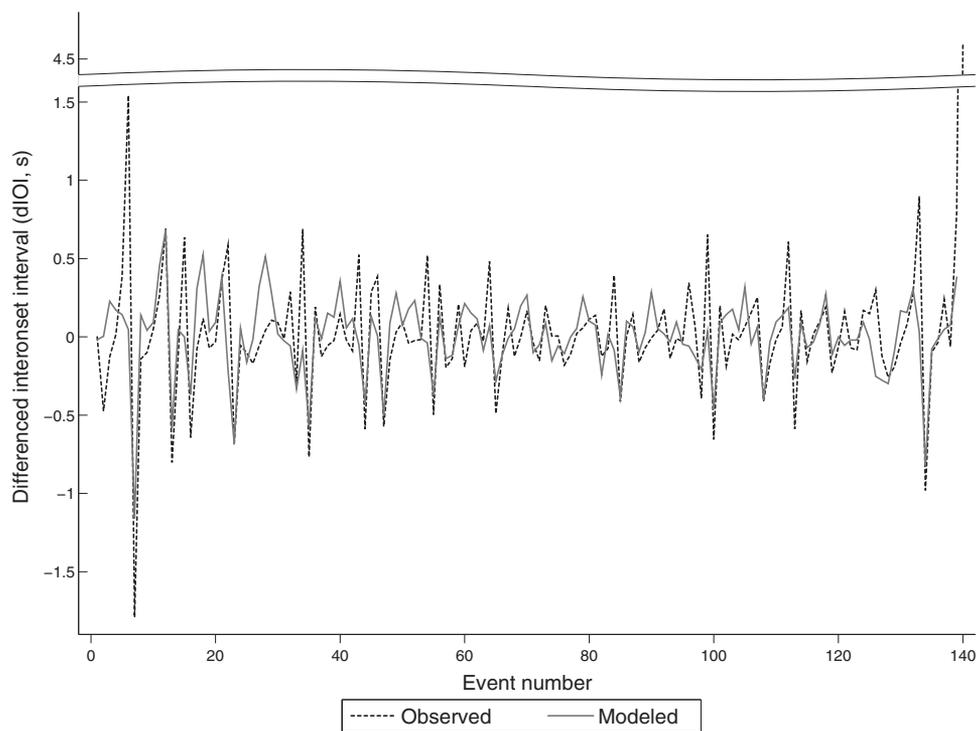
although the value should not be viewed as identical in nature to a correlation between sets of independent variables. The correlations were similar for the default and optimized models, suggesting little advantage of the optimized set of features. Figure 4 compares the dIOI to its prediction from the IC of the optimized model, showing the good fit achieved.

While the dEntropy was not a significant predictor of dIOI, we hypothesized that it might be a stronger predictor of interperformer variability, and this idea was supported in the best model we obtained. Thus we determined a time series for the CV of the IOIs across performers (abbreviated CV–IOI) and developed a model of it. CV–IOI was stationary and AR, and hence could be modeled without differencing. By a similar analytical process to that just

outlined, we found that in our best model CV–IOI was predicted jointly by Lead 1 of IC for the optimized model and Leads 1 and 2 of its entropy, with AR terms. This model (also shown in Table 1) provided satisfactory residuals, and the modeled-observed correlation was 0.52 (see Figure 5 for the comparison of modeled and observed time series). Once again, the model was substantially better than the best purely AR model, suggesting that our predictors did indeed capture variation in the data.

## Discussion

As expected, changes in IC predicted performance timing. Specifically, performers appeared to slow down in anticipation of



*Figure 4.* Observed and modeled differenced interonset intervals for each event of the *Prélude non mesuré No. 7*. Note the discontinuity in the vertical axis.

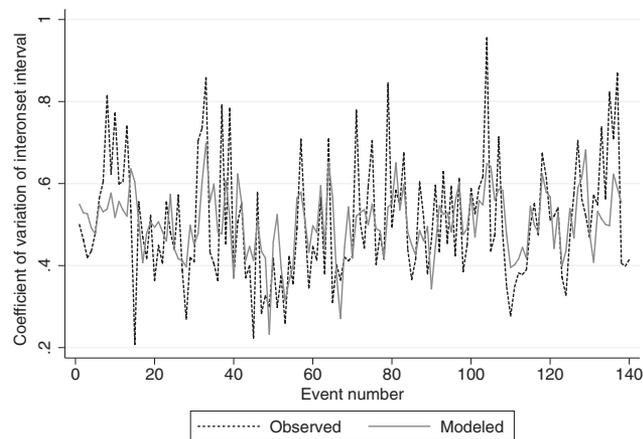


Figure 5. Observed and modeled coefficients of variation of interonset interval for each event of the *Prélude non mesuré No. 7*.

unexpected notes (notes characterized by a high IC). No differences appeared between the default and optimized versions of the IDyOM model. Entropy was not a significant predictor.

Turning to interperformer variability, both IC and entropy were predictors. Therefore, performance timing appeared to be more variable among performers around points of high uncertainty and unexpectedness in the musical structure. Modeling of two individual performances that represented extremes of duration (not shown) revealed that they could be quite well fit by the standard models shown in Table 1, supporting our interpretations.

An alternative explanation for the link we observed between variations in expressive timing in performance and entropy would be that performers are constrained to slow down at more challenging points in the piece (due to technical difficulties) and that these challenging points likely correspond to points of greater uncertainty in the piece. However, this explanation seems implausible given that all performers were professional harpsichordists who received the score of the piece several weeks prior to the recording session, and that the error rates were very low. Furthermore, previous research has shown that, among expert populations such as professional performers, trade-offs between timing and accuracy only become apparent when individuals perform at the limit of their abilities and under strict temporal constraints (Pfordresher, Palmer, & Jungers, 2007). Our subjective impression from listening to the excerpts was that the performers were mostly or fully in control of the timing strategies they employed, so it is very unlikely that the performers would have been constrained to slow down at certain points due to technical difficulties (besides, the piece itself would not be considered to be particularly challenging for professional harpsichordists, as author R.T. Dean, himself a professional pianist, can confirm).

In Experiment 2, we examined the impact of both structural predictability and performance timing on the experience of musical tension in harpsichordists, other musicians, as well as nonmusicians listening to the performances analyzed in Experiment 1. To do so, we used IC and entropy as model-based predictors of IOI and musical tension (using our information-dynamic model of melodic structure), and IOI as an expressive, performance-derived predictor of musical tension.

## Experiment 2: Perceived Tension

### Method

**Participants and stimuli.** Ten harpsichordists, 20 nonharpsichordist musicians, and 20 nonmusicians were invited to listen to recordings of the performances collected in Experiment 1. The harpsichordist listeners (eight male and two female) had an average age of 31.6 years (range: 23–38 years). None of these harpsichordists had participated in Experiment 1. The musician participants (12 male and eight female) had a minimum of 2 years in undergraduate music studies and had an average age of 26.7 years (range: 20–47 years). The nonmusician participants (12 male and eight female) had less than 2 years of musical training in early childhood and had an average age of 26.6 years (range: 20–48 years). Listeners received monetary compensation for participation.

Only one recording from each performer who participated in Experiment 1 was used, yielding a set of 12 stimuli, each from a different performer. Performances were chosen subjectively as the best one (of two for each performer) in terms of recording quality and fewest performance errors.

**Procedure.** All nonharpsichordist musicians and nonmusicians, as well as five harpsichordists, were tested at McGill University, Montreal, Canada. Because of the difficulty of recruiting a sufficient number of Montreal-based harpsichordists who had not participated in Experiment 1, an additional five harpsichordists were tested at Goldsmiths, London, England. Aside from some minor technical detail differences between the two experimental setups, the experimental procedure was identical.

In both cities, the experiment took place in a double-walled, IAC sound-isolation chamber; the sound booth models used were Model 120-act3 in Montreal and Model 1200-A in London (IAC Acoustics, Bronx, NY). In Montreal, participants performed the experiment on a Macintosh G5 computer running OS 10.6.8 (Apple Computer, Cupertino, CA). Sounds were amplified through a Grace Design m904 monitor (Grace Digital Audio, San Diego, CA) and presented over Sennheiser HD280 Pro headphones (Sennheiser Electronic GmbH, Wedemark, Germany). In London, participants performed the experiment on a 15-in. MacBook Pro running OS 10.6.3 (Apple Computer, Cupertino, CA). Sounds were played through an Edirol UA-4FX USB external sound card (Roland Corp., Shizuoka, Japan) and presented over Sennheiser HD202 headphones (Sennheiser Electronic GmbH, Wedemark, Germany).

All participants passed an audiogram to ensure that they had normal hearing. The musical excerpts were presented at a comfortable listening level kept constant for all participants. Participants used a slider to continuously rate the perceived musical tension of each performance. The slider was equipped with an elastic rubber band to provide feedback through physical resistance as it was moved toward higher tension ratings. The slider was connected to an AtoMIC Pro analog-to-MIDI converter (Ircam-Centre Pompidou, Paris, France). The computer interface, programmed in the PsiExp environment (Smith, 1995), provided instructions on the screen and allowed participants to advance through the trials by clicking on an on-screen button with the mouse. Participants were instructed to hold the slider box on the desk or on their lap and to move the slider forward (away from themselves) to indicate higher perceived tension and backward (toward themselves) to indicate lower perceived tension. They were

asked to use the full range of the slider for at least one performance during the whole experiment.

The experiment was divided into two parts. During the first part (two practice trials), the experimenter was present to answer questions. The practice performances were taken from the set of 12 performances not used in Experiment 2 (see the Participants and Stimuli section). Participants were instructed that the two practice trials were representative of the range of performances in the main part of the experiment. Data from the practice trials were not analyzed. The definition of musical tension was left open due to the wide range of participants involved with the experiment. The main experimental block consisted of one performance by each performer, interspersed with 15 s of nature sounds as a short break. The order of performances was randomized for each participant. Once the experiment was completed, participants filled out a questionnaire. The entire experiment lasted approximately 30 min.

**Data analysis.** For each participant, a log file recorded the tension slider ratings continuously over time for each performance. Due to the force exerted by the elastic band when stretched at maximum, a block was inserted into the slider, resulting in MIDI ratings ranging between 54 and 127, which were then range normalized by participant. We used the matched score of the MIDI data of the performances to establish a correspondence between MIDI note onset events and score. The tension values were averaged over each score event.

Statistical models predicting perceived tension on the basis of the performance timing (IOI) and IDyOM parameters (IC and entropy) only made use of the IOI data from the 12 performances heard by participants in Experiment 2 (i.e., within-performer average IOI time series were not used here, in contrast to Experiment 1). Here, we also used VAR to provide the Granger causality test (Granger, 1969, 2004), which yields a principled analysis of the directionality of the relations between factors that might be mutually influential in a pair or group of time series (taking account of the autocorrelations). In other words, Granger causality analysis shows which time series (e.g., IOI, IC) provide information that can predict future values of other time series, which was the main purpose of our analyses. Granger causality analyses were conducted to determine whether IOI and/or IDyOM parameter time series were significant predictors of future values of the perceived tension profiles. Other aspects of the TSA followed the procedure described in Experiment 1.

## Results

As shown in Figure 6, the profiles of perceived tension for the musician and nonmusician groups differed somewhat, whereas harpsichordists' profiles were very similar to those of the nonharpsichordist musicians. Thus, we first modeled the possible impact of the IOI and IDyOM parameters on the perceived tension profiles of each participant group independently.

The tension profiles showed strong autocorrelation. The first differenced tension profile was stationary in the case of musicians and harpsichordists and could be modeled satisfactorily. VAR analyses showed that IDyOM parameters had predictive influence, and that they provided a fair model of the tension profile. However, when they were taken together with the IOI, the results showed that only the IOI was required for the optimal model of the tension profile, which was significantly improved. In other words, IC and entropy were not

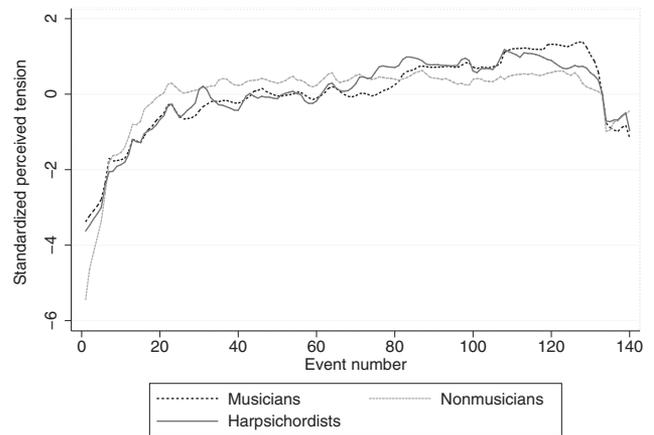


Figure 6. Standardized ( $z$ -scored) perceived tension for the harpsichordists, nonharpsichordist musicians, and nonmusicians for each event of the *Prélude non mesuré* No. 7.

required predictors. Moreover, unlike IC or entropy, IOI was a Granger causal predictor of the tension profile.

Because the VAR analysis, and in particular Granger causality, provides a direct comparison of the IOI and information-theoretic series as predictors of perceived tension, it is worth illustrating this further, taking as an example the case of the musicians as auditors. We optimized a VAR model (of the first differenced variables) in which all four time series (tension, IC, entropy, and IOI) were considered as endogenous. That is, in the model, they may influence each other from a statistical perspective, and a joint model of all four variables is constructed. This was the most conservative way of evaluating the model, because IC and entropy were computed directly from the score of the *Prélude* and were thus fixed (exogenous), and could not be influenced by the other components. We obtained a reasonable model of dTension, with  $R^2$  equal to .40 for the fit. The Granger analysis showed that the only Granger causal predictor of tension was dIOI ( $p < .01$ );  $p$  values for dIC and dEntropy were not significant ( $ps > .31$ ). Furthermore, the coefficients for the dIC and dEntropy series were also not significant at any lag included in the model ( $ps > .29$ ). Thus, dIOI was clearly the only useful predictor. Very similar results were obtained for the harpsichordists as auditors, given the slightly different overall model parameters (see below). On the basis of these and other VAR analyses, we undertook ARMAX models comparable to those presented in Experiment 1.

Table 2 summarizes the best ARMAX time-series modeling results for perceived tension, giving white noise residuals. Figure 7 shows the average differenced profile for tension as perceived by nonharpsichordist musicians and its prediction on the basis of the best ARMAX model. Although harpsichordists showed similar perceived tension profiles to nonharpsichordist musicians, the best model for their perception involved the Lead 1 ("F") of dIOI rather than its Lag 1 (see Table 2). This lead dIOI parameter could not be substituted in models by the IC features that predicted dIOI (as shown in Table 1).

In the case of the nonmusicians' tension profile, two differencing steps were required to reach stationarity. As with harpsichordists and nonharpsichordist musicians, only IOI time series showed Granger causality for the tension profiles of nonmusicians. Correspondingly, only IOI series had a significant influence, besides

Table 2

*Time-Series Analysis of Tension Ratings (dTension) Using Information Content and Entropy (From Two IDyOM Models, Using Default and Optimized Viewpoints) and Performance Timing (dIOI) as Predictors*

Modeled parameter	Predictors considered (IDyOM model in parentheses)	Components of the optimal model (coefficients for the external predictors in parentheses)	BIC	Correlation coefficients (modeled vs. observed values)
dTension Musicians	IC and entropy (default), dIOI	L1.dIOI (+.75), AR(1, 5, 6)	521.0	.65
dTension Musicians	IC and entropy (optimized), dIOI	L1.dIOI (+.75), AR(1, 5, 6)	521.0	.65
d2Tension Nonmusicians	IC and entropy (optimized), dIOI	L1.d2IOI (+.04), F1.d2IOI (-.13), AR(1-5, 8)	513.4	.61
dTension Harpsichordists	IC and entropy (optimized), dIOI	F(1, 2).dIOI (-.65, +.31), AR(1, 2)	540.4	.54

*Note.* The chosen data series were stationary; all models listed showed satisfactory white noise residuals. For example, L1.dIOI (.75) AR(1, 5, 6) (row 1) means: lead one of the (first) differenced IOI series, derived from the optimized model, had a coefficient of +.75. See the Method section in Experiment 1 for further details of the time-series analysis. Note that although a range of parameters is considered as possible predictors (column 2), a smaller selection remains in the optimal model (as shown in column 3). Thus, for example, the models of rows 1-2 are identical. AR(n) = autoregressive error terms; BIC = Bayesian information criterion; CV = coefficient of variation; Fn = lead n; IC = information content; IDyOM = Information Dynamics of Music; IOI = interonset interval; Ln = lag n; MA(n) = moving average error term lag n, where n represents the number of events.

AR, in ARMAX analysis of stationarized tension profiles of non-musicians (see Figure 8).

## Discussion

Our analysis of the perceived tension data indicated that it is expressive performance timing (analyzed here in terms of IOI time series), rather than information-dynamic parameters such as IC and entropy, that directly influences perceived tension for all groups of listeners. Indeed, neither IC nor entropy were Granger-causal nor required predictors in statistical models predicting perceived tension, whereas IOI (more specifically, its differenced values) entered all models. Notably, there was no significant additional influence of IC on the tension values once the effect of expressive timing was taken into account, and only expressive timing was a Granger-causal predictor of perceived tension. However, as shown in Experiment 1, IC and entropy predicted performance timing itself. Similarly, it was confirmed in Experiment 2 that IC and entropy can predict perceived tension, although they were not required predictors, and did not account for any variance above and beyond what is already accounted for by the IOI values (which they predict as shown in Experiment 1). Thus, we may most readily interpret Experiment 2 as showing that the

information parameters (extracted from the score as conceived by the composer) influence expressive performance timing (the realization by performers), which in turn influences perceived tension (the reception by listeners), embodying the tripartite communication model of Kendall and Carterette presented in the Introduction.

Table 2 also revealed an interesting difference between the harpsichordists' tension ratings and those of the nonmusicians and musicians. Indeed, unlike nonmusicians and nonharpsichordist musicians, harpsichordists' ratings were predicted only by imminent events (reflected in the F or "lead" term in the TSA) rather than by immediately preceding events (represented by the L or "lag" term in Table 2). This suggests that harpsichordists' tension ratings do in fact anticipate upcoming events, perhaps due to their greater familiarity with the musical style.

## General Discussion

To our knowledge, this is the first large-scale study to combine an analysis of expressive strategies in performance, melodic expectations (as predicted by a computational model), and listeners' ratings of musical tension to characterize the relations among these elements of musical communication. Results showed that, in a

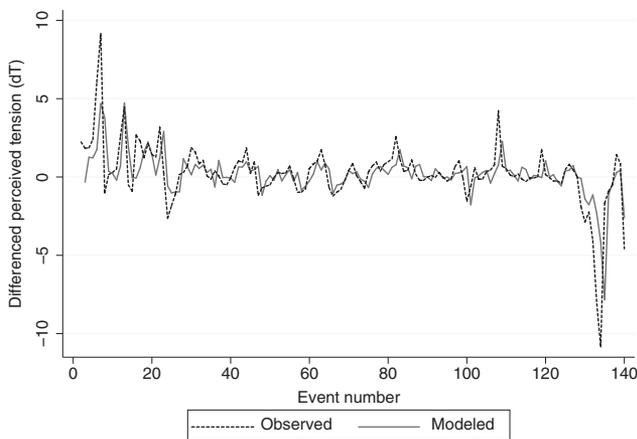


Figure 7. Observed and modeled differenced perceived tension for each event of the *Prélude non mesuré No. 7* for the nonharpsichordist musicians.

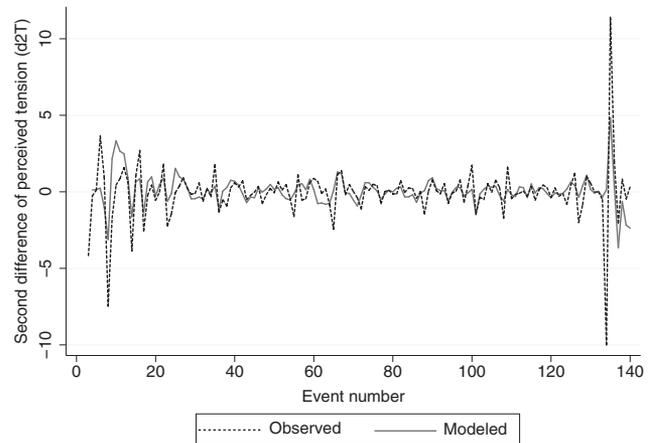


Figure 8. Observed and modeled second difference of perceived tension for each event of the *Prélude non mesuré No. 7* for the nonmusicians.

semi-improvisatory musical genre such as the unmeasured prelude, the composed melodic structure, as measured by information-dynamic parameters, has the capacity to predict local tempo variations in performance. These, in turn, predict the perception of tension by listeners, including harpsichordists, nonharpsichordist musicians, and nonmusicians, in broadly similar ways across groups. More generally, our findings indicated that, when the constraints on expressive timing in performance introduced by notated meter and note duration are removed, the link between melodic structure and expressive timing becomes statistically salient, thus bringing to the fore the close interconnection between the dimensions of pitch and time in music.

Our results thus substantiate the proposed interactions in the tripartite relation suggested by Kendall and Carterette (1990): the cognitive processes involved in the communication between composer (via the notated score) and performer, and also the interactive communicative effect of the melodic expectations implied by the score, combined with the expressive timing in performance, on the listener. Furthermore our findings illustrate that these relations act in a cumulative manner, in which one chain of influence (composer → performer) indirectly affects another (performer → listener).

Our choice of a piece containing a single unaccompanied melodic line, played on an instrument with a narrow dynamic range such as the harpsichord, allowed us to avoid the complexities related to polyphony and harmony and variations in sound intensity, without compromising ecological validity. To be sure, our results do not necessarily indicate that melodic structure is the main predictor of musical tension. However, because sound intensity and roughness were presumably very similar for all performances (besides the limited dynamic range of harpsichord music, all recordings were made on the same instrument and under the same acoustical conditions), the experiments presented here allow us to understand more precisely the influence of pitch predictability and performance timing on perceived tension. In future work, it will also be interesting to assess whether sound intensity profiles, even in harpsichord performances, have a bearing on the perception of affect, and whether they themselves are sufficiently varied across the piece to reflect any influence of the flux of IC. Whether or not variations in intensity in harpsichord performances are perceived by listeners, the fact that harpsichordists appear to display individual velocity profiles (Gingras et al., 2013) and that they tend to play an emphasized voice with higher velocity (Gingras, Asselin, Goodchild, & McAdams, 2009) suggests that velocity profiles do play a role, however marginal, in expressive harpsichord performance.

The piece selected for this study exhibited a relatively simple, high-level structure, and we predicted that influences of this structural layout can also be detected in the performance timings.<sup>3</sup> Furthermore, we surmised that such structural influences generalize in more complex pieces to other aspects of performance and perception, as we have observed in pieces from a range of musical styles (Bailes & Dean, 2012). Once we are able to deal with structures of sufficient complexity, it may be possible to enhance information-theoretic properties of the event structures by larger scale musicological ones (such as the “development” or “recapitulation” structures of classical music sonatas).

Previous research using the IDyOM model has demonstrated that IC accurately accounts for pitch expectations generated while listening to melodies (Pearce, 2005; Pearce et al., 2010; Omigie et al., 2012), whereas entropy predicts listeners’ uncertainty about the note

that will follow in a given context (Hansen & Pearce, 2014). These previous results have suggested that perceptual processing of pitch sequences can be accurately modeled as a cognitive process of probabilistic prediction based on implicit statistical learning through exposure to music. The present results extend this picture for the first time to expressive performance and to emotional experience during listening. Specifically, the results confirmed that IC (or unexpectedness) is a good predictor of performance timing and indicate that the entropy (or uncertainty) may also be influential. The results also exhibited a relation between IC and tension experienced by listeners, although there was no additional variance in the tension ratings accounted for by information-theoretic variables after accounting for performance timing (which was itself predicted by IC and entropy). Current work is extending IDyOM to the analysis of other aspects of music, such as harmony (multiple pitches notated with simultaneous timing in the score, absent in this piece) and meter (entirely lacking from the current piece), and their interactions. This approach will be used to understand the influence of information-theoretic properties of more complex stimuli, including metrical and harmonic structure (in addition to melodic structure, examined here) on performance and perception.

It is interesting to place this piece, which is an example of a genre that codifies prior improvisations, within the current activities that develop from it: free improvisation. This immediately points to the fact that our study could be extended by applying a similar methodology to free musical improvisations, in which not only are the meter and duration not given (as with the piece we used), but also the melodic structure is elaborated by the performer. A series of related studies on improvisation has suggested that this approach could be fruitful (Dean & Bailes, 2014; Dean, Bailes, & Drummond, 2014). Thus, we surmise that the melodic structure of an improvisation should also predict timing in performance and perceived tension, especially for music without regular meter such as free jazz and free improvisation, although this is commonly done in circumstances permitting wide ranges of acoustic intensity, which are expected also to be influential.

In conditions of experimentally controlled event timings, it may be theoretically possible to disentangle the relative contribution of the pitch structure and expressive timing to perceived tension by exposing listeners to artificial isochronous performances (with all the notes having the same duration) and comparing the results with those obtained through actual performances of the same pieces. However, a stimulus with equal-duration pitches (isochronous rhythm) would cause listeners to perceive a regular accent pattern (Povel, 1984), which is incompatible with the specific constraints of this essentially unmeasured musical genre. To isolate the contribution of the pitch structure, one might also envision using artificial performances in which all notes have the same pitch height. However, Lerdahl and Krumhansl (2007) have argued that musical expectations (closely related to musical tension) are essentially undefined for repeated pitches, suggesting that IC and entropy may not be appropriate measures for a stimulus with repeated identical pitches. Despite the above-mentioned drawbacks associated with isochronous perfor-

<sup>3</sup> A companion article compares performers’ structural analyses of the *Prélude* with their performance timings and the resulting perceived tension profiles, both on a local level and considering the large-scale structure of the piece (Goodchild et al., in press).

mances or performances with monotonic pitches, studies using such stimuli might yield useful insights that help us determine whether our findings can be generalized to other musical genres. Moreover, a thorough experimental investigation of the impact of systematically controlled temporal deviations on perceived tension would be necessary to proceed beyond our demonstration of Granger causality (a strong directional statistical statement) to a claim of reductive strong causality. Such experimental manipulations are beyond our scope here, but should be pursued in future research.

This work has clear implications for music composition, performance, and listening, the three poles of the model proposed by Kendall and Carterette. For composition, it confirms that systematic organization of pitch (and presumably other structures) measurably influences the performers' interpretive choices as well as the listeners' emotional responses, and can be used to shape these responses with a degree of predictability. For performance, it suggests that implicit or explicit awareness of the musical structure—of both the broader stylistic conventions to which a piece relates and of the piece itself—is likely to have a significant impact on the choice of expressive strategies and on the resulting expressive impact on the audience. Regarding the aesthetic appreciation of music by listeners, it again suggests that implicit knowledge of such statistical/structural features influences the tension experienced during musical listening and that familiarity with a musical idiom can transform that influence.

We have focused on music (in fact, a specific piece of music with particular properties) as a domain in which the relations between performance and perception can be investigated and related to the structure of the stimulus in a controlled way, without compromising ecological validity. However, the results have implications for human perception and performance beyond music. The role of prediction has been established in several areas of psychology, such as language processing (Cristià et al., 2011; DeLong et al., 2005; Hale, 2006; Levy, 2008; Saffran, 2003), visual perception (Bar, 2007; Bubic et al., 2010; Egner et al., 2010), and motor sequencing (Wolpert & Flanagan, 2001). Furthermore, probabilistic approaches have been proposed to model the cognitive processes involved in acquiring and generating predictions in these domains (e.g., Friston, 2010; Levy, 2008; Perruchet & Pacton, 2006). Taking language as an example, the present findings are consistent with existing results that have shown the information-theoretic structure of a sentence predicts aspects of performance such as word reading times (see, e.g., Levy, 2008), and suggests that these, in turn, may predict emotional responses, at least for certain kinds of linguistic production such as poetry. The same may be true of visual domains involving structured sequences of human movement, such as dance or silent film. These hypotheses should be tested in future research. In so doing, the framework we have used here provides a rigorous methodology for identifying quantitative dynamic relations between perception, performance, and the structure of the stimulus.

## References

- Bailes, F., & Dean, R. T. (2012). Comparative time series analysis of perceptual responses to electroacoustic music. *Music Perception, 29*, 359–375. <http://dx.doi.org/10.1525/mp.2012.29.4.359>
- Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences, 11*, 280–289. <http://dx.doi.org/10.1016/j.tics.2007.05.005>
- Bharucha, J. J., & Stoeckig, K. (1986). Reaction time and musical expectancy: Priming of chords. *Journal of Experimental Psychology: Human Perception and Performance, 12*, 403–410. <http://dx.doi.org/10.1037/0096-1523.12.4.403>
- Bharucha, J. J., & Stoeckig, K. (1987). Priming of chords: Spreading activation or overlapping frequency spectra? *Perception & Psychophysics, 41*, 519–524. <http://dx.doi.org/10.3758/BF03210486>
- Bubic, A., von Cramon, D. Y., & Schubotz, R. I. (2010). Prediction, cognition and the brain. *Frontiers in Human Neuroscience, 4*, 25.
- Carlsen, J. C. (1981). Some factors which influence melodic expectancy. *Psychomusicology: Music, Mind & Brain, 1*, 12–29. <http://dx.doi.org/10.1037/h0094276>
- Castellano, M. A., Bharucha, J. J., & Krumhansl, C. L. (1984). Tonal hierarchies in the music of north India. *Journal of Experimental Psychology: General, 113*, 394–412. <http://dx.doi.org/10.1037/0096-3445.113.3.394>
- Clarke, E. F. (1985). Structure and expression in rhythmic performance. In P. Howell, I. Cross, & R. West (Eds.), *Musical structure and cognition* (pp. 209–236). London, England: Academic Press.
- Clarke, E. F. (1988). Generative principles in music performance. In J. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation and composition* (pp. 1–26). New York, NY: Oxford University Press.
- Clarke, E. F. (1989). The perception of expressive timing in music. *Psychological Research, 51*, 2–9. <http://dx.doi.org/10.1007/BF00309269>
- Conklin, D., & Witten, I. H. (1995). Multiple viewpoint systems for music prediction. *Journal of New Music Research, 24*, 51–73. <http://dx.doi.org/10.1080/09298219508570672>
- Cooper, W. E., & Danly, M. (1981). Segmental and temporal aspects of utterance-final lengthening. *Phonetica, 38*, 106–115. <http://dx.doi.org/10.1159/000260017>
- Cristià, A., McGuire, G. L., Seidl, A., & Francis, A. L. (2011). Effects of the distribution of acoustic cues on infants' perception of sibilants. *Journal of Phonetics, 39*, 388–402. <http://dx.doi.org/10.1016/j.wocn.2011.02.004>
- Dean, R. T., & Bailes, F. (2010). Time series analysis as a method to examine acoustical influences on real-time perception of music. *Empirical Musicology Review, 5*, 152–175.
- Dean, R. T., & Bailes, F. (2014). Cognitive processes in musical improvisation. In G. E. Lewis & B. Piekut (Eds.), *The Oxford handbook of critical improvisation studies* (Vol. 1). New York, NY: Oxford University Press. Advance online publication. <http://dx.doi.org/10.1093/oxfordhb/9780195370935.013.007>
- Dean, R. T., Bailes, F., & Drummond, J. (2014). Generative structures in improvisation: Computational segmentation of keyboard performances. *Journal of New Music Research, 43*, 224–236. <http://dx.doi.org/10.1080/09298215.2013.859710>
- Dean, R. T., Bailes, F., & Dunsmuir, W. T. M. (2014a). Time series analysis of real-time music perception: Approaches to the assessment of individual and expertise differences in perception of expressed affect. *Journal of Mathematics and Music: Mathematical and Computational Approaches to Music Theory, Analysis, Composition and Performance, 8*, 183–205. <http://dx.doi.org/10.1080/17459737.2014.928752>
- Dean, R. T., Bailes, F., & Dunsmuir, W. T. M. (2014b). Shared and distinct mechanisms of individual and expertise-group perception of expressed arousal in four works. *Journal of Mathematics and Music: Mathematical and Computational Approaches to Music Theory, Analysis, Composition and Performance, 8*, 207–223. <http://dx.doi.org/10.1080/17459737.2014.928753>
- DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*, 1117–1121. <http://dx.doi.org/10.1038/nn1504>
- Eerola, T. (2004). Data-driven influences on melodic expectancy: Continuations in North Sami Yoiks rated by South African traditional healers. In S. Lipscomb, R. Ashley, R. O. Gjerdingen, & P. Webster (eds.), *Proceedings of the Eighth International Conference on Music Perception and Cognition* (pp. 83–87). Adelaide, Australia: Causal Productions.

- Egermann, H., Pearce, M. T., Wiggins, G. A., & McAdams, S. (2013). Probabilistic models of expectation violation predict psychophysiological emotional responses to live concert music. *Cognitive, Affective & Behavioral Neuroscience*, *13*, 533–553. <http://dx.doi.org/10.3758/s13415-013-0161-y>
- Egner, T., Monti, J. M., & Summerfield, C. (2010). Expectation and surprise determine neural population responses in the ventral visual stream. *The Journal of Neuroscience*, *30*, 16601–16608. <http://dx.doi.org/10.1523/JNEUROSCI.2770-10.2010>
- Farbood, M. M. (2012). A parametric, temporal model of musical tension. *Music Perception*, *29*, 387–428. <http://dx.doi.org/10.1525/mp.2012.29.4.387>
- Fletcher, N. H. (1977). Analysis of the design and performance of harpsichords. *Acustica*, *37*, 139–147.
- Fredrickson, W. E. (1995). A comparison of perceived musical tension and aesthetic response. *Psychology of Music*, *23*, 81–87. <http://dx.doi.org/10.1177/0305735695231006>
- Fredrickson, W. E. (2000). Perception of tension in music: Musicians versus nonmusicians. *Journal of Music Therapy*, *37*, 40–50. <http://dx.doi.org/10.1093/jmt/37.1.40>
- Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*, 127–138. <http://dx.doi.org/10.1038/nrn2787>
- Gabrielsson, A., & Juslin, P. N. (1996). Emotional expression in music performance: Between the performer's intention and the listener's experience. *Psychology of Music*, *24*, 68–91. <http://dx.doi.org/10.1177/0305735696241007>
- Gingras, B. (2008). *Expressive strategies and performer-listener communication in organ performance* (Doctoral dissertation, McGill University).
- Gingras, B., Asselin, P. Y., Goodchild, M., & McAdams, S. (2009). Communicating voice emphasis in harpsichord performance. In J. Louhivuori, T. Eerola, S. Saarikallio, T. Himberg, & P. Eerola (Eds.), *Proceedings of the 7th Triennial Conference of the European Society for the Cognitive Sciences of Music* (pp. 154–158). Jyväskylä, Finland: University of Jyväskylä.
- Gingras, B., Asselin, P.-Y., & McAdams, S. (2013). Individuality in harpsichord performance: Disentangling performer- and piece-specific influences on interpretive choices. *Frontiers in Psychology*, *4*, 895. <http://dx.doi.org/10.3389/fpsyg.2013.00895>
- Gingras, B., & McAdams, S. (2011). Improved score-performance matching using both structural and temporal information from MIDI recordings. *Journal of New Music Research*, *40*, 43–57. <http://dx.doi.org/10.1080/09298215.2010.545422>
- Goebl, W. (2001). Melody lead in piano performance: Expressive device or artifact? *The Journal of the Acoustical Society of America*, *110*, 563–572. <http://dx.doi.org/10.1121/1.1376133>
- Goodchild, M., Gingras, B., & McAdams, S. (in press). Analysis, performance, and tension perception of an unmeasured prelude for harpsichord. *Music Perception*.
- Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, *37*, 424–438. <http://dx.doi.org/10.2307/1912791>
- Granger, C. W. J. (2004). Time series analysis, cointegration, and applications. *The American Economic Review*, *94*, 421–425. <http://dx.doi.org/10.1257/0002828041464669>
- Hale, J. (2006). Uncertainty about the rest of the sentence. *Cognitive Science*, *30*, 643–672. [http://dx.doi.org/10.1207/s15516709cog0000\\_64](http://dx.doi.org/10.1207/s15516709cog0000_64)
- Hansen, N. C., & Pearce, M. T. (2014). Predictive uncertainty in auditory sequence processing. *Frontiers in Psychology*, *5*, 1052. <http://dx.doi.org/10.3389/fpsyg.2014.01052>
- Hanslick, E., & Cohen, G. (1854). *Vom musikalisch-Schönen* [On the beautiful in music]. Darmstadt, Germany: Wissenschaftliche Buchgesellschaft.
- Huron, D. B. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: MIT Press.
- Ilie, G., & Thompson, W. F. (2006). A comparison of acoustic cues in music and speech for three dimensions of affect. *Music Perception*, *23*, 319–330. <http://dx.doi.org/10.1525/mp.2006.23.4.319>
- Juslin, P. N., & Laukka, P. (2004). Expression, perception, and induction of musical emotions: A review and a questionnaire study of everyday listening. *Journal of New Music Research*, *33*, 217–238. <http://dx.doi.org/10.1080/0929821042000317813>
- Juslin, P. N., & Timmers, R. (2011). Expression and communication of emotion in music performance. In P. N. Juslin & J. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 453–489). New York, NY: Oxford University Press.
- Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, *31*, 559–575. <http://dx.doi.org/10.1017/S0140525X08005293>
- Kendall, R. A., & Carterette, E. C. (1990). The communication of musical expression. *Music Perception*, *8*, 129–163. <http://dx.doi.org/10.2307/40285493>
- Kessler, E. J., Hansen, C., & Shepard, R. N. (1984). Tonal schemata in the perception of music in Bali and in the West. *Music Perception*, *2*, 131–165. <http://dx.doi.org/10.2307/40285289>
- Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Friederici, A. D. (2004). Music, language and meaning: Brain signatures of semantic processing. *Nature Neuroscience*, *7*, 302–307. <http://dx.doi.org/10.1038/nn1197>
- Krumhansl, C. L. (1990). Tonal hierarchies and rare intervals in music cognition. *Music Perception*, *7*, 309–324. <http://dx.doi.org/10.2307/40285467>
- Krumhansl, C. L. (1996). A perceptual analysis of Mozart's Piano Sonata K. 282: Segmentation, tension, and musical ideas. *Music Perception*, *13*, 401–432. <http://dx.doi.org/10.2307/40286177>
- Krumhansl, C. L. (2002). Music: A link between cognition and emotion. *Current Directions in Psychological Science*, *11*, 45–50. <http://dx.doi.org/10.1111/1467-8721.00165>
- Krumhansl, C. L., Louhivuori, J., Toiviainen, P., Järvinen, T., & Eerola, T. (1999). Melodic expectation in Finnish spiritual folk hymns: Convergence of statistical, behavioral, and computational approaches. *Music Perception*, *17*, 151–195. <http://dx.doi.org/10.2307/40285890>
- Krumhansl, C. L., Toivanen, P., Eerola, T., Toiviainen, P., Järvinen, T., & Louhivuori, J. (2000). Cross-cultural music cognition: Cognitive methodology applied to North Sami Yoiks. *Cognition*, *76*, 13–58. [http://dx.doi.org/10.1016/S0010-0277\(00\)00068-8](http://dx.doi.org/10.1016/S0010-0277(00)00068-8)
- Large, E. W., Palmer, C., & Pollack, J. B. (1995). Reduced memory representations for music. *Cognitive Science*, *19*, 53–96. [http://dx.doi.org/10.1207/s15516709cog1901\\_2](http://dx.doi.org/10.1207/s15516709cog1901_2)
- Lerdahl, F. (2001). *Tonal pitch space*. New York, NY: Oxford University Press.
- Lerdahl, F., & Krumhansl, C. L. (2007). Modeling tonal tension. *Music Perception*, *24*, 329–366. <http://dx.doi.org/10.1525/mp.2007.24.4.329>
- Levitin, D. J., & Tirovolas, A. K. (2009). Current advances in the cognitive neuroscience of music. *Annals of the New York Academy of Sciences*, *1156*, 211–231. <http://dx.doi.org/10.1111/j.1749-6632.2009.04417.x>
- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*, 1126–1177. <http://dx.doi.org/10.1016/j.cognition.2007.05.006>
- Madsen, C. K. (1997). Emotional response to music as measured by the two-dimensional CRDI. *Journal of Music Therapy*, *34*, 187–199. <http://dx.doi.org/10.1093/jmt/34.3.187>
- Madsen, C. K., Brittin, R. V., & Capperella-Sheldon, D. A. (1993). An empirical method for measuring the aesthetic experience to music. *Journal of Research in Music Education*, *41*, 57–69. <http://dx.doi.org/10.2307/3345480>
- Madsen, C. K., & Fredrickson, W. E. (1993). The experience of musical tension: A replication of Nielsen's research using the continuous re-

- sponse digital interface. *Journal of Music Therapy*, 30, 46–63. <http://dx.doi.org/10.1093/jmt/30.1.46>
- Manning, C. D., & Schütze, H. (1999). *Foundations of statistical natural language processing*. Cambridge, MA: MIT Press.
- McAdams, S., Vines, B., Vieillard, S., Smith, B., & Reynolds, R. (2004). Influences of large-scale form on continuous ratings in response to a contemporary piece in a live concert setting. *Music Perception*, 22, 297–350. <http://dx.doi.org/10.1525/mp.2004.22.2.297>
- Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago, IL: University of Chicago Press.
- Narmour, E. (1990). *The analysis and cognition of basic melodic structures: The implication–realization model*. Chicago, IL: University of Chicago Press.
- Omigie, D., Pearce, M. T., & Stewart, L. (2012). Tracking of pitch probabilities in congenital amusia. *Neuropsychologia*, 50, 1483–1493. <http://dx.doi.org/10.1016/j.neuropsychologia.2012.02.034>
- Omigie, D., Pearce, M. T., Williamson, V. J., & Stewart, L. (2013). Electrophysiological correlates of melodic processing in congenital amusia. *Neuropsychologia*, 51, 1749–1762. <http://dx.doi.org/10.1016/j.neuropsychologia.2013.05.010>
- Oram, N., & Cuddy, L. L. (1995). Responsiveness of Western adults to pitch-distributional information in melodic sequences. *Psychological Research*, 57, 103–118.
- Palmer, C. (1989). Mapping musical thought to musical performance. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 331–346. <http://dx.doi.org/10.1037/0096-1523.15.2.331>
- Palmer, C. (1996). On the assignment of structure in music performance. *Music Perception*, 14, 23–56. <http://dx.doi.org/10.2307/40285708>
- Panksepp, J., & Bernatzky, G. (2002). Emotional sounds and the brain: The neuro-affective foundations of musical appreciation. *Behavioural Processes*, 60, 133–155. [http://dx.doi.org/10.1016/S0376-6357\(02\)00080-3](http://dx.doi.org/10.1016/S0376-6357(02)00080-3)
- Pearce, M. T. (2005). *The construction and evaluation of statistical models of melodic structure in music perception and composition* (Doctoral dissertation, City University, London). Retrieved from <http://openaccess.city.ac.uk/8459/>
- Pearce, M. T., Conklin, D., & Wiggins, G. (2005). Methods for combining statistical models of music. In U. K. Wiil (Ed.), *Computer music modeling and retrieval* (pp. 295–312). Berlin, Germany: Springer. [http://dx.doi.org/10.1007/978-3-540-31807-1\\_22](http://dx.doi.org/10.1007/978-3-540-31807-1_22)
- Pearce, M. T., Ruiz, M. H., Kapasi, S., Wiggins, G. A., & Bhattacharya, J. (2010). Unsupervised statistical learning underpins computational, behavioural, and neural manifestations of musical expectation. *NeuroImage*, 50, 302–313. <http://dx.doi.org/10.1016/j.neuroimage.2009.12.019>
- Pearce, M. T., & Wiggins, G. (2004). Improved methods for statistical modelling of monophonic music. *Journal of New Music Research*, 33, 367–385. <http://dx.doi.org/10.1080/0929821052000343840>
- Pearce, M. T., & Wiggins, G. A. (2006). Expectation in melody: The influence of context and learning. *Music Perception*, 23, 377–405. <http://dx.doi.org/10.1525/mp.2006.23.5.377>
- Pearce, M. T., & Wiggins, G. A. (2012). Auditory expectation: The information dynamics of music perception and cognition. *Topics in Cognitive Science*, 4, 625–652. <http://dx.doi.org/10.1111/j.1756-8765.2012.01214.x>
- Penttinen, H. (2006). On the dynamics of the harpsichord and its synthesis. In V. Verfaillie (Ed.), *Proceedings of the 9th International Conference on Digital Audio Effects (DAFx-06)* (pp. 115–120). Montreal, Canada.
- Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, 10, 233–238. <http://dx.doi.org/10.1016/j.tics.2006.03.006>
- Pfordresher, P. Q., Palmer, C., & Jungers, M. K. (2007). Speed, accuracy, and serial order in sequence production. *Cognitive Science*, 31, 63–98. <http://dx.doi.org/10.1080/03640210709336985>
- Povel, D. J. (1984). A theoretical framework for rhythm perception. *Psychological Research*, 45, 315–337. <http://dx.doi.org/10.1007/BF00309709>
- Repp, B. H. (1990). Patterns of expressive timing in performances of a Beethoven minuet by nineteen famous pianists. *The Journal of the Acoustical Society of America*, 88, 622–641. <http://dx.doi.org/10.1121/1.399766>
- Repp, B. H. (1992). Diversity and commonality in music performance: An analysis of timing microstructure in Schumann’s “Träumerei.” *Journal of the Acoustical Society of America*, 92, 2546–2568. <http://dx.doi.org/10.1121/1.404425>
- Repp, B. H. (1995). Expressive timing in Schumann’s “Träumerei”: An analysis of performances by graduate student pianists. *Journal of the Acoustical Society of America*, 98, 2413–2427. <http://dx.doi.org/10.1121/1.413276>
- Repp, B. H. (1996). The art of inaccuracy: Why pianists’ errors are difficult to hear. *Music Perception*, 14, 161–183. <http://dx.doi.org/10.2307/40285716>
- Saffran, J. R. (2003). Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science*, 12, 110–114. <http://dx.doi.org/10.1111/1467-8721.01243>
- Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, 70, 27–52. [http://dx.doi.org/10.1016/S0010-0277\(98\)00075-4](http://dx.doi.org/10.1016/S0010-0277(98)00075-4)
- Schellenberg, E. G. (1997). Simplifying the implication-realization model of melodic expectancy. *Music Perception*, 14, 295–318. <http://dx.doi.org/10.2307/40285723>
- Schellenberg, E. G., Adachi, M., Purdy, K. T., & McKinnon, M. C. (2002). Expectancy in melody: Tests of children and adults. *Journal of Experimental Psychology: General*, 131, 511–537. <http://dx.doi.org/10.1037/0096-3445.131.4.511>
- Schwarz, G. E. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6, 461–464. <http://dx.doi.org/10.1214/aos/1176344136>
- Shaffer, L. H., & Todd, N. P. (1987). The interpretive component in musical performance. In A. Gabriellson (Ed.), *Action and perception in rhythm and music* (No. 55, pp. 139–152). Stockholm, Sweden; Royal Swedish Academy of Music.
- Shepard, R. N. (1982). Structural representations of musical pitch. In D. Deutsch (Ed.), *The psychology of music* (pp. 343–390). Orlando, FL: Academic Press. <http://dx.doi.org/10.1016/B978-0-12-213562-0.50015-2>
- Sloboda, J. A. (2010). Music in everyday life: The role of emotions. In P. N. Juslin & J. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 493–514). New York, NY: Oxford University Press.
- Smith, B. K. (1995). *PsiExp: An environment for psychoacoustic experimentation using the IRCAM musical workstation*. Paper presented at the Society for Music Perception and Cognition, University of California, Berkeley, CA.
- Steinbeis, N., Koelsch, S., & Sloboda, J. A. (2006). The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. *Journal of Cognitive Neuroscience*, 18, 1380–1393. <http://dx.doi.org/10.1162/jocn.2006.18.8.1380>
- Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review*, 107, 885–913. <http://dx.doi.org/10.1037/0033-295X.107.4.885>
- Tillmann, B., Bigand, E., & Pineau, M. (1998). Effects of global and local contexts on harmonic expectancy. *Music Perception*, 16, 99–117. <http://dx.doi.org/10.2307/40285780>
- Tilney, C. (Ed.). (1991). *The art of the French unmeasured prelude* (Vol. 2; H. Lerch, trans.). London, England: Schott Music.
- Todd, N. (1985). A model of expressive timing in tonal music. *Music Perception*, 3, 33–57. <http://dx.doi.org/10.2307/40285321>

Toiviainen, P., & Krumhansl, C. L. (2003). Measuring and modeling real-time responses to music: The dynamics of tonality induction. *Perception, 32*, 741–766. <http://dx.doi.org/10.1068/p3312>

Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current Biology, 11*, R729–R732. [http://dx.doi.org/10.1016/S0960-9822\(01\)00432-8](http://dx.doi.org/10.1016/S0960-9822(01)00432-8)

## Appendix

### Viewpoints Used With IDyOM

In this research, we used two versions of IDyOM (Information Dynamics of Music; Pearce, 2005), one using the *default* viewpoints and the other using the *optimized* viewpoints. The default model uses a single viewpoint linking pitch interval (the pitch difference in semitones between a note and its immediate predecessor; denoted below by *cpint*) with scale degree (the difference in semitones of a note from the tonic of the key; denoted below by *cpintfref*). The optimized version uses a set of

viewpoints optimized through hill-climbing to maximize the predictability of the musical piece. See the main text for a detailed description of IDyOM (in the section “The Composition: Musical Structure, Predictability, and Expectation”).

Table A1 gives the full set of viewpoints submitted to the optimization procedure, which also considered all possible pairwise links between these individual viewpoints. The viewpoints selected for the optimized model are shown in Table A2.

Table A1

*The Viewpoints Submitted to the Selection Procedure for the Optimized Model*

Viewpoint	Description
cpitch	Chromatic pitch (MIDI pitch number)
cpitch-class	Octave equivalent pitch class (chroma)
tessitura	Three values: whether a note is between 66 (G#4) and 74 (D5), above, or below this range
cpint	Pitch interval in semitones
cpint-size	Absolute size of pitch interval
cpcint	Pitch interval class (cpint mod 12)
cpcint-size	Absolute size of cpcint
contour	Pitch contour (–1, 0, 1)
newcontour	Boolean: whether or not contour is the same as the previous contour
cpintfip	Pitch interval from the first note in the piece
cpintfref	Chromatic scale degree
inscale	Boolean: whether or not the note is in the scale

Table A2

*The Viewpoints Selected for the Optimized Model*

Viewpoint added	Mean information content
cpint-size	4.849
tessitura × inscale	4.430
cpcint-size × cpintfref	4.296
cpitch × cpint-size	4.149
cpitch-class × cpintfref	4.080
cpitch × cpintfip	4.055
cpintfref × inscale	4.054
tessitura × cpcint	4.044
cpcint	4.037

Received June 14, 2014

Revision received August 7, 2015

Accepted August 7, 2015 ■