Effects of Pitch and Timing Expectancy on Musical Emotion

Sarah A. Sauvé and Aminah Sayed
Queen Mary University of London

Roger T. Dean
Western Sydney University

Marcus T. Pearce
Queen Mary University of London

Pitch and timing information work hand in hand to create a coherent piece of music, but what happens when this information goes against the norm? Relationships between musical expectancy and emotional responses were investigated in a study conducted with 40 participants: 20 musicians and 20 nonmusicians. Participants took part in one of two behavioral paradigms measuring continuous expectancy or emotional responses (arousal and valence) while listening to folk melodies that exhibited either high or low pitch predictability and high- or low- onset predictability. The causal influence of pitch predictability was investigated in an additional condition in which pitch was artificially manipulated and a comparison was conducted between original and manipulated forms; the dynamic correlative influence of pitch and timing information and its perception on emotional change during listening was evaluated using cross-sectional time series analysis. The results indicate that pitch and onset predictability are consistent predictors of perceived expectancy and emotional response, with onset carrying more weight than pitch. In addition, musicians and nonmusicians do not differ in their responses possibly due to shared cultural background and knowledge. The results demonstrate in a controlled lab-based setting a precise, quantitative relationship between the predictability of musical structure, expectation, and emotional response.

Keywords: emotion, expectation, time series, information content, IDyOM

Music is capable of inducing powerful physiological and psychological emotional states (Bittman et al., 2013; Castillo-Pérez, Gómez-Pérez, Velasco, Pérez-Campos, & Mayoral, 2010; Habibi & Damasio, 2014). For example, the practice of music therapy stemmed from the observation that music can have a positive emotional effect (Khalfa, Bella, Roy, Peretz, & Lupien, 2003; Pelletier, 2004). However, many studies of emotion induction by music have simply investigated which emotions are induced rather than the psychological mechanisms that account for why these emotions occur (Juslin & Västfjäll, 2008). The present research aims to address this omission by examining a theorized psychological mechanism of musical emotion induction in isolation. Although factors such as personality, age, and gender have an influence (Rentfrow & Gosling, 2003), we focus here on the properties of music that are involved in emotion induction.

Although there is general consensus that music can elicit emotional responses (see Juslin & Sloboda, 2011 for an extensive review), why and how it does so is less clear. Juslin and colleagues (Juslin, Liljeström, Västfjäll, & Lundqvist, 2011) described eight potential psychological mechanisms that might explain how emotions are induced through music: (a) brain stem reflexes, (b) evaluative conditioning, (c) emotional contagion, (d) visual imagery, (e) episodic memory, (f) musical expectancy, (g) rhythmic entrainment, and (h) cognitive appraisal. Hearing a sudden loud or dissonant event causes a change in arousal (brain stem reflex), whereas a piece repetitively paired with a positive, or negative,

---

situation will create a positive, or negative, emotional reaction (evaluative conditioning). Emotional contagion is the induction of emotion through the mimicry of behavioral or vocal expression of emotion and is reflected in musical structure; for example, shorter durations and ascending pitch contours tend to reflect happiness, whereas longer durations and descending pitch contours communicate sadness. Visual imagery refers to the mental imagery evoked by the music, which can have positive or negative affect. The pairing between a sound and a past event can trigger the emotion related to that event when the sound is subsequently heard (episodic memory). Rhythmic entrainment refers to the induction of emotion through the proprioceptive feedback of internal body entrainment (i.e., heart rate) to the music and, finally, cognitive appraisal refers to the evaluation of music in the context of goals or plans of the listener. The present study focuses on musical expectancy while controlling for all other potential mechanisms proposed in the abovementioned text.

Meyer (1956) argued that emotion is generated through musical listening because listeners actively generate predictions reflecting what they expect to hear next (Huron, 2006). Unexpected events are surprising and are associated with an increase in tension, whereas expected events are associated with the resolution of tension (Gingras et al., 2016). According to this account, surprising events generally evoke high arousal and low valence (Egermann, Pearce, Wiggins, & McAdams, 2013; Koelsch, Fritz, & Schlaug, 2008; Russell, 2003; Steinbeis, Koelsch, & Sloboda, 2006). Although the arousal response to increased tension is fairly consistent, listeners familiar with a piece of music can come to appreciate an event that has low expectancy through an appraisal mechanism, resulting in a high valence response (Huron, 2006). This apparent contradiction highlights the importance of isolating the psychological mechanisms behind musical emotional induction.

There are also different influences on musical expectation (Huron, 2006). Schematic influences reflect general stylistic patterns acquired through extensive musical listening to many pieces of music, whereas veridical influences reflect specific knowledge of a familiar piece of music. Dynamic influences reflect dynamic learning of structure within an unfamiliar piece of music (e.g., recognizing a repeated motif). Listening to new, unfamiliar music in a familiar style engages schematic and dynamic mechanisms, the former reflecting long-term learning over years of musical exposure and the latter, short-term learning within an individual piece of music. Both these long- and short-term mechanisms can be simulated as a process of statistical learning and probabilistic generation of expectations (Pearce, 2005). Furthermore, these may be different for musicians and nonmusicians due to extensive exposure and training in a particular style (Juslin & Västfjäll, 2008).

We now consider the properties of musical events for which expectations are generated. Prominent among such properties are the pitch and timing of notes, and we consider each in turn. Music theorists have described musical styles as structurally organized, reflecting well-formalized rules that constitute a kind of grammatical syntax (Lerdahl & Jackendoff, 1983). In the tradition of Western tonal music, compositions usually follow these rules by adhering to a tonal structure, and enculturated listeners are able to identify when a piece of music infringes tonal rules based on their exposure in everyday music listening (Carlsen, 1981; Krumhansl, Louhivuori, Toiviainen, Järvinen, & Eerola, 1999; Trainor & Trehub, 1992). Two kinds of models have been developed to explain listeners' pitch expectations: first, models that include static rules and, second, models that focus on learning. An influential example of a rule-based model is the implication-realization model, developed by Eugene Narmour (1991), which includes rules defining the expectedness of the final note in a sequence of three notes, in which the first pair of notes forms the implicative interval and the second pair of notes, the realized interval. The size and direction of the implicative interval set up expectations of varying strengths for the realized interval. Although the original implication-realization model contained five bottom-up rules of melodic implication, Schellenberg (1996) reduced the five bottom-up rules of the implication-realization model to two: *pitch proximity* and *pitch reversal*. For example, according to the rules of pitch reversal, a small interval implies another small interval in the same direction, whereas a large interval implies a subsequent small interval in the opposite direction. Such patterns reflect actual patterns in existing Western music (Huron, 2006), suggesting the possibility that listeners might learn these patterns through experience.

Statistical learning is a powerful tool for explaining the acquisition of pitch expectations in music, in which common sequential patterns are learned through incidental exposure (Huron, 2006; Pearce, 2005; Saffran, Johnson, Aslin, & Newport, 1999), making them more predictable to the exposed listener. For example, a perfect cadence is found at the end of the vast majority of Western classical music, in which movement from dominant to tonic is the strongest form of closure possible in this style. Through repeated exposure to this pattern, a dominant penultimate chord strongly implies a tonic chord for an enculturated listener (Huron, 2006). Information Dynamics of Music (IDyOM; Pearce, 2005) is a computational model of auditory expectation that harnesses the power of statistical learning. It learns the frequencies of variable-order musical patterns from a large corpus of music (via the long-term model) and from the current piece of music being processed (via the short-term model or short-term memory) in an unsupervised manner and generates probabilistic predictions about the properties of the next note in a melody, given the preceding melodic context. IDyOM is a multiple-viewpoint model capable of learning patterns from pitch- and time-derived note properties (source viewpoints) to predict relevant note properties (target viewpoints). These viewpoints can be use-defined or selected through optimization. The information content (negative log probability; IC) of an event, given the model, reflects the unexpectedness of the event in context. Low IC corresponds to high expectedness, whereas high IC corresponds to low expectedness.

Temporal regularities are also learned through exposure (Cirelli, Spinelli, Nozaradan, & Trainor, 2016; Hannon, Soley, & Ullal, 2012; Hannon & Trehub, 2005a, 2005b). Western music is dominated by beat patterns in divisions of two, and to a lesser extent, divisions of three, and compositional devices much as syncopation and hemiola (three beats in the time of two) are used to manipulate the listener's temporal expectations. The dynamic attending theory (Jones & Boltz, 1989) posits that entrainment to a beat results in attentional focus being directed at time intervals in which a beat is expected to occur, such that longer entrainment times result in stronger predictions and more focused attention. This was supported using a pitch discrimination task in which participants performed better on pitch discrimination when target pitches fell on expected time points

as a result of entrainment to a series of isochronous distractor tones (Jones, Moynihan, MacKenzie, & Puente, 2002; though see Bauer, Jaeger, Thorne, Bendixen, & Debener, 2015 for conflicting evidence). We propose that temporal rules can also be explained by statistical learning as implemented in IDyOM (Pearce, 2005). In the same way as pitch, and various other musical surface structures, onset and interonset interval (IOI) can be predicted by IDyOM, as it learns from a given corpus and a given piece. This is equivalent to estimating a distribution over possible future onset times, given the preceding sequences of events. Because pitch and temporal structures generate distinct expectancies (Prince, Thompson, & Schmuckler, 2009), we explore the influence of each as a potential emotional inducer using both correlational and causal methods (while allowing for the possibility of interactions between pitch and timing).

Musical expectancy as a mechanism for the induction of emotion in listeners has been studied in an ecological setting: Egermann et al. (2013) asked 50 participants to attend a live concert during which six flute pieces were played. These pieces spanned various musical styles and levels of pitch expectancy. Three kinds of measurement were made: subjective responses (i.e., the arousal levels or the ratings of musical expectancy that changed continuously throughout the piece), expressive responses (using video and facial electromyography), and peripheral arousal measured by skin conductance, heart rate, respiration, and blood volume pulse. IDyOM (Pearce, 2005) was used to analyze pitch patterns of the music and predict where listeners would experience low expectancy. Results suggested that expectancy had a modest influence on emotional responses, in which high IC segments led to higher arousal and lower valence ratings, as well as increases in skin conductance and decreases in heart rate, as compared with low IC segments, whereas no event-related changes were found in respiration rate or facial electromyography measures; however, this study was conducted in an ecologically valid, thus noncontrolled environment where participants could have focused on something other than the music. For example, visual aspects of performance are highly important to emotional engagement in live music settings (Thompson, Graham, & Russo, 2005; Vines, Krumhansl, Wanderley, & Levitin, 2006). Furthermore, other potential emotion-inducing mechanisms, as proposed by Juslin and Västfjäll (2008), were not explicitly controlled for, and effects of temporal expectancy on emotional responses were not considered.

The current study is designed to investigate pitch and temporal musical expectancy in a restricted environment that controlled for many other potential emotional mechanisms (Juslin & Västfjäll, 2008). Brain stem reflexes are controlled for by maintaining equal tempo, intensity, and timbre across all musical excerpts. Evaluative conditioning and episodic memory are controlled for by presenting unfamiliar musical excerpts, so that expectation ratings and emotional reactions are not confounded by previous experience with the music. Potential effects of emotional contagion are analyzed in the analysis by including pitch and IOI as predictors of subjective ratings and pitch and IOI predictability (i.e., higher mean pitch and shorter IOI could result in higher arousal and valence ratings regardless of expectancy). Irrelevant visual imagery cannot be categorically avoided but the rating tasks are expected to require enough cognitive load to render it unlikely. Furthermore, to the extent that visual imagery is variable between individuals, averaging across participants should remove its influence. The absence of a strong, driving beat and the relatively short duration of the musical excerpts makes deep, emotion-inducing rhythmic entrainment highly unlikely. Finally, all participants are listening to these musical excerpts in the context of an experiment, with any other goal or motive being highly unlikely, thus minimizing the relevance of the cognitive appraisal mechanism.

This research aims to address three questions. First, does the predictability of pitch and timing (as simulated by IDyOM) have an effect on listeners' expectations and emotional state, and can we causally influence this effect with explicit manipulation of the stimuli? We hypothesize that the degree of musical expectancy for pitch (based on pitch interval) and temporal (based on IOI) structures, as predicted objectively by the information content provided by IDyOM, will have an effect on emotion as measured by the arousal-valence model (Russell, 2003). According to Russell (2003), unexpected events will invoke negative valence and cause an increase in arousal, and expected events will invoke positive valence and decreased arousal. We do not expect appraisal to affect this initial reaction, as we are collecting ratings in real time. We also hypothesize that when both pitch and timing are either expected or unexpected, the emotional response will be more extreme than in conditions of mixed expectedness. Furthermore, direct manipulation of pitch expectancy while keeping temporal expectancy and all other musical features constant is expected to produce the predicted changes in ratings (i.e., transforming unexpected pitches to expected pitches will decrease unexpectedness and arousal, and increase valence ratings).

Second, how does pitch and timing predictability combine to influence expectation and emotion? Though the combination of pitch and timing in music perception has been a research interest for decades (Boltz, 1999; Duane, 2013; Jones, Boltz, & Kidd, 1982; Palmer & Krumhansl, 1987a; Prince et al., 2009), no clear conclusions can be drawn, as findings regarding this question have low agreement and seem highly dependent on the choice of stimuli, participants, and paradigm. For example, although Prince et al. (2009) suggest that pitch is more salient, results from Duane's (2013) corpus analysis suggest that timing is the most reliable predictor of streaming. Although the present study uses monophonic melodies, it could be argued that if salience is linked to complexity (Prince et al., 2009), for melodies in which pitch or timing are highly predictable (low complexity), the predictable feature will be less salient than its unpredictable counterpart because it requires less "processing power" and therefore less attention. For melodies in which pitch and timing are relatively equally predictable or unpredictable, their relative importance currently remains unknown.

Finally, is there a difference in the effect of pitch and timing predictability on expectation and emotional responses between musicians and nonmusicians? The effect of musical training will be evaluated by comparing the responses of musicians and nonmusicians, with the expectation that musicians will have higher expectation ratings and more extreme emotional responses to pitch and timing violations due to training (Strait, Kraus, Skoe, &

Ashley, 2009), in which Western musical patterns are more familiar, resulting in violations of these patterns eliciting stronger responses.

Method

Participants

Forty participants (22 female and 18 male; age range 14–54 years) were recruited from universities, secondary school, and colleges for this experiment: 20 were musicians (*Mean* 3.6 years of musical training; range 1–12 years) and 20 were nonmusicians (0 years of musical training). Ethical approval was obtained from the Queen Mary Research Ethics Committee, QMREC1078.

Stimuli

The stimuli consisted of 32 pieces of music in Musical Instrument Digital Interface (MIDI) format rendered to audio using a piano timbre: 16 original melodies and 16 artificially manipulated melodies. Original melodies were divided into the following four categories of predictability: predictable pitch and predictable onset (PP), predictable pitch and unpredictable onset (PU), unpredictable pitch and predictable onset (UP) and unpredictable pitch and unpredictable onset (UU). The artificial melodies were created by changing the pitch predictability of each melody, so that PP became aUP, UU became aPU, PU became aUU, and UP became aPP, where *a* denotes artificial. All melodies were presented at the same intensity, which was held constant for the duration of all melodies.

Original melodies. The 16 original melodies were selected from a group of nine data sets, totaling 1,834 melodies (Table 1;

Table 1
Details of the Data Sets Used in Stimulus Selection

Data set	Description	Number of melodies	Mean events/ composition
2	Chorale soprano melodies harmonized by J.S. Bach	100	46.93
3	Alsatian folk songs from the Essen Folk Song Collection	91	49.40
4	Yugoslavian folk songs from the Essen Folk Song Collection	119	22.61
5	Swiss folk songs from the Essen Folk Song Collection	93	49.31
6	Austrian folk songs from the Essen Folk Song Collection	104	51.01
10	German folk songs from the Essen Folk Song Collection: ballad	687	40.24
15	German folk songs from the Essen Folk Song Collection: kinder	213	39.40
18	British folk song fragments used in the experiments of Schellenberg (1996)	8	18.25
23	Irish folk songs encoded by Daiman Sagrillo	62	78.50

Figure 1. Excerpts from one melody from each of the four different types of experimental stimuli. Patterns or notes of interest are marked with a bracket or an arrow, respectively. Melody PP is predictable in both pitch and time; an exact repetition in both dimensions can be seen, marked by a square bracket. Melody PU is predictable in pitch but unpredictable in time; long notes in general and the rhythmic switch in the last measure specifically contribute to low predictability. Melody UP is unpredictable in pitch but predictable in time, with large leaps (marked by arrow) and regular note durations. Melody UU is unpredictable in both pitch and time; a leap is surprising after such repetitive unison, and the bracketed rhythmic excerpt is a hemiola (here 3 notes in the time of 2).

Figure 1), all from Western musical cultures to avoid potential cultural influences on expectancy ratings (Hannon & Trehub, 2005b; Palmer & Krumhansl, 1990). All nine data sets were analyzed by IDyOM for target viewpoints pitch and onset with source viewpoints pitch interval and scale degree (linked), and IOI, respectively. Both short-term and long-term models were engaged; the long-term l model was trained on three data sets of Western music, described in Table 2. There was no resampling within the test data sets.

The 1,834 melodies were divided into four categories based on high or low pitch or onset IC. Melodies were considered predictable if they had a lower IC than the mean IC of all samples and unpredictable if the IC was greater than the mean IC of all samples. Four melodies from each category were selected as the most or least predictable by finding maximum and minimum IC values as appropriate for the category; these are the original 16 melodies. Melodies in the PP, PU, UP, and UU categories had mean pitch IC values ranging from 1.37 to 1.85, 2.22 to 2.43, 2.83 to 5.24, and 2.61 to 2.78, respectively; mean onset IC values ranging from .80 to .92, 2.49 to 4.34, 1.13 to 1.32 and 4.20 to 4.39, respectively; mean MIDI pitch values (i.e., 69 = A = 440 Hz) ranging from 66.85 to 70.17, 66.05 to 70.23, 68.67 to 72.76, and 64.40 to 71.63, respectively; and mean IOI values ranging from 12.71 to 21.28, 21.41 to

Table 2
Details of the Training Set Used to Train IDyOM

Data set	Description	Number of melodies	Mean events/ composition
0	Songs and ballads from Nova Scotia, Canada	152	56.26
1	Chorale melodies harmonized by J.S. Bach	185	49.88
7	German folk songs	566	58.46

69.81, 13.84 to 21.69, and 21.53 to 64.00, respectively, where a quarter note equals 24, an eighth note, 12, a half note, 48, and so forth Notice that categories with unpredictable onset have higher average IOI values; this potential influence is discussed in the text below (Table A1 in Appendix A).

Artificial melodies. The 16 artificial melodies were created as follows. For each original melody, the notes with the highest (for PP and PU) or lowest (for UP and UU) IC were selected for replacement. The notes were replaced with another note from the same melody that shares the same preceding note as the original note in that melody. If several instances of such a note pair existed, the associated IC values were averaged. If several such replacement notes existed, the one with the lowest (for UP and UU) or highest (for PP and PU) IC was selected to replace the original note. Where no such replacement existed, the key of the melody was estimated using the Krumhansl–Schmuckler key-finding algorithm (Krumhansl, 2001) using key profiles updated by Temperley (1999), and the replacement was selected as the scale degree with the highest (for UP and UU) or lowest (for PP and PU) tonal stability. All notes labeled as having extremely high or low IC were replaced by a pitch with a less extreme IC. An example of a melody from each category can be seen in Figure 1.

Melodies in the aPP, aPU, aUP, and aUU categories had mean pitch IC values ranging from 3.49 to 5.50, 4.20 to 4.56, 4.13 to 6.59, and 2.79 to 3.80, respectively, and mean MIDI pitch values ranging from 64.88 to 69.80, 67.05 to 73.18, 64.05 to 67.76, and 66.78 to 72.89, respectively. Mean onset IC and mean raw IOI values were unchanged from the corresponding original stimulus-predictability category (e.g., aPP has the same mean IOI IC and IOI values as UP). Figure 2 illustrates the mean information content of all 32 melodies.

Procedure

Participants were semirandomly allocated to one of four (between-subjects) conditions: They were either a musician or a nonmusician and, within these groups, randomly assigned to rate either expectancy or emotion (arousal and valence). The experiment was run on a software constructed in-house, and on a Samsung Galaxy Ace S5830 (3.5 in. in diameter; running Android 2.3.6). Participants listened through standard Apple headphones and were tested individually in a closed room. The information sheet was presented and informed consent gathered; detailed step-by-step instructions were then presented to the participants. Regardless of condition, there was a mandatory practice session: Participants heard two melodies and answered the questions appropriate to the condition they were assigned to (either expectancy rating or arousal and valence rating). Participants could also adjust the volume to a comfortable setting during the practice session. Once the practice session was completed, the experimental app was loaded. Participants entered a unique identity (ID) number provided by the experimenter and responded to a short musical background questionnaire. Participants then heard the 32 musical excerpts (mean duration was 18.34 s) presented in a random order without pause or repeat and performed the appropriate ratings by continuously moving a finger on the screen. Those in the expectancy rating condition reported expectancy on a 7-point
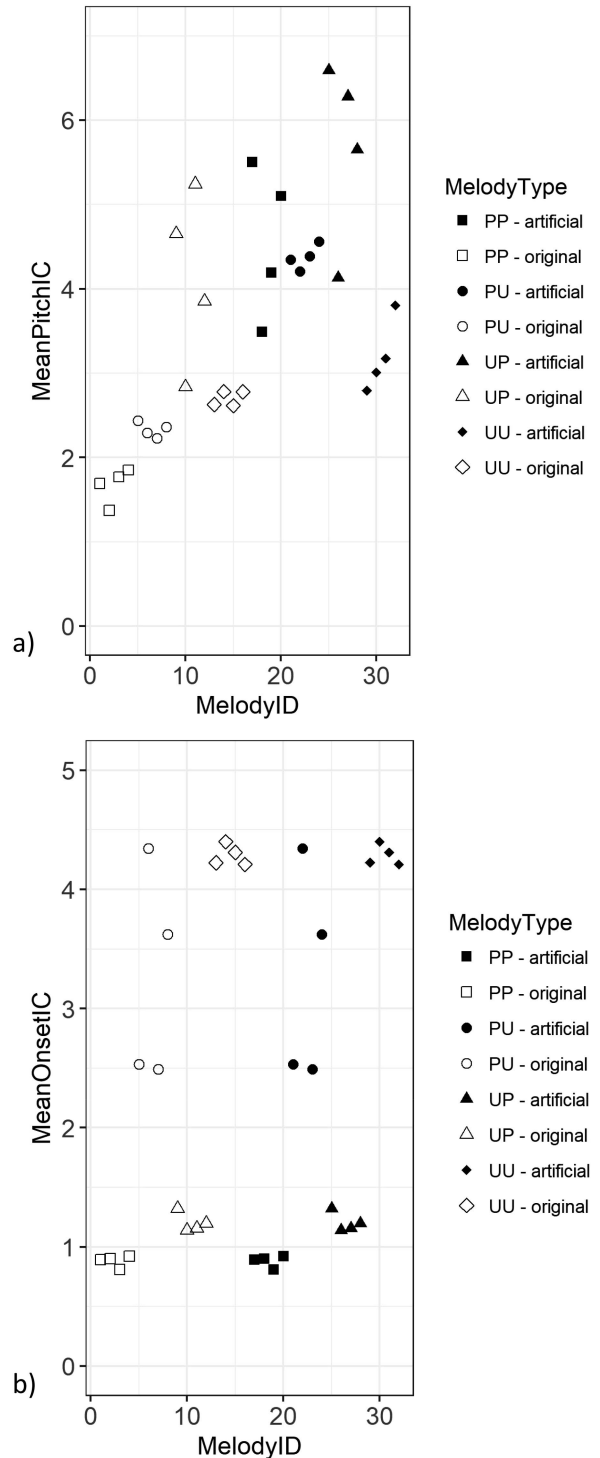

Figure 2. Mean (a) pitch information content and (b) onset IC of each melody plotted by stimulus predictability and modification, in which original melodies are symbolized by empty symbols and artificial melodies by full symbols.

integer Likert scale, where 1 was *very expected* and 7 was *very unexpected*. Those in the arousal/valence condition rated induced arousal (vertical) and valence (horizontal) on a two-dimensional arousal/valence illustration (Russell, 2003). Responses, in integers, were collected at a 5 Hz sample rate (200 ms; Khalfa, Isabelle, Jean-Pierre, & Manon, 2002). The rating systems used were Expectancy: 1–7 (*expected–unexpected*); Arousal: 0–230 (*calm–stimulating*); Valence: 0–230 (*unpleasant–pleasant*).

Data Collection

Due to the large number of variables included in this analysis, we describe them each here for clarity. First, we describe the dependent variables: expectancy, arousal, and valence ratings. Ratings, on scales described previously, were collected at a rate of 5 Hz, or 200 ms, making it the variable with the smallest temporal resolution. Thus, all other variables were interpolated to match this resolution.

We then begin describing the independent variables with those that we are not explicitly manipulating: time, pitch, IOI, and musicianship. Time is measured in steps of 200 ms, the sampling rate of the data acquisition software. For each point in time, for each melody, and each participant, a *data point*, a value for *pitch*, *IOI*, *musical training*, *stimulus predictability*, *stimulus modification*, *pitch IC* and *onset IC* is assigned, along with *Melody ID* and *Participant ID*. Because pitch (interpreted here in MIDI numbers) does not change every 200 ms and IOI (in ms) is longer than 200 ms in these folk songs (or in Western music in general), their values were interpolated to match the participant ratings' sampling rate of 5 Hz, so that each point in time has a pitch and IOI value, and values were simply duplicated. Finally, the musical training variable had a value of 0 or 1, depending on whether the participant had no musical training or any musical training, respectively.

Next, we describe manipulated variables: stimulus predictability, stimulus modification, pitch IC, and onset IC. For each data point, the variable stimulus predictability was given a value of 1 if it belonged to the category PP, 2 if it belonged to the category PU, 3 for UP, and 4 for UU, regardless of whether these are original or artificial melodies. Similarly, the variable stimulus modification was given a value of 0 if the melody was original or 1 if the melody was artificial. Finally, pitch IC and onset IC, as calculated by IDyOM, were interpolated in the same way as pitch and IOI to match the participant ratings' 5 Hz sampling rate. These are the only variables whose values are not integers.

Though the duplication of data points due to interpolation is taken into account by modeling discontinuous time in the case of melody-level analysis and autoregression in the case of cross-sectional time series analysis (CSTSA), a data set without interpolation was also created to corroborate any findings using interpolated data, in which ratings were averaged according to the rate of change of each melody's events. In other words, each melody event was assigned one rating value, with 1 associated MIDI pitch, IOI, pitch IC, and onset IC.

Statistical Analysis

For each type of rating (expectancy, arousal, and valence) two kinds of analysis are performed: first, a melody-level analysis, in which the time series for each melody are averaged across partic-

ipants separately for musicians and nonmusicians, leading to approximately 9,600 data points, and temporal position is a discontinuous factor and, second, a cross-sectional time-series analysis of the continuous ratings given by each participant throughout each melody, leading to approximately 96,000 data points. In the melody-level analysis, for each melody, a mean expectancy rating was calculated at every time point across the musician and nonmusician groups (10 responses per group). Linear multiple regression modeling was used to evaluate the effect of *musical training* (musician or nonmusician), *stimulus modification* (original or artificial), and *stimulus predictability* (predictable/unpredictable pitch/onset) by in turn comparing a model containing each single predictor with a model containing only an intercept using a log likelihood test. Two additional predictors, *pitch predictability* and *onset predictability*, were derived from stimulus predictability to examine the interaction between these two subcomponents: Melodies were coded as having either predictable or unpredictable pitch or onset. Although musical training, stimulus modification, pitch predictability, and onset predictability were simple binary factors, stimulus predictability contained four levels labeled PP, PU, UP, and UU. Apart from that, between pitch predictability and onset predictability, interactions were not considered due to the difficulty of interpretation in such a complex model. Following these log likelihood comparisons, two *global* linear multiple regression models, containing all the above predictors of interest (one containing stimulus predictability and the other containing pitch predictability and onset predictability) plus *time*, *pitch*, and *IOI* to parse out any potential effects of time and to analyze potential effects of musical contagion, were evaluated to confirm results.

For the analysis of continuous ratings throughout each melody, we used CSTSA, similar to Dean, Bailes, and Dunsmuir (2014a), to evaluate the predictive impacts of *pitch IC*, *onset IC*, *stimulus predictability* (predictable/unpredictable), *stimulus modification* (none/artificial), *musical training*, and individual differences modeled by random effects on participants' ratings of expectedness, arousal, and valence. CSTSA takes account of the autoregressive characteristic of music and the continuous responses of the participants. Pitch IC and onset IC predictors were both scaled to values between 0 and 1, to allow for direct comparison of model coefficients in analysis. A predictor of combined *pitch and onset IC* was also tested, replacing the individual *pitch IC* and *onset IC* predictors. In practice, CSTSA is a mixed-effects multiple linear regression model, here fitted with maximum design-driven random effects (Barr, Levy, Scheepers, & Tily, 2013) and fixed effects to account for autocorrelation (lags of endogenous variables, i.e., ratings, denoted by P), and exogenous influence (i.e., information content and its lags, denoted by L). Here, each lag window represents 200 ms, so that a predictor named *L1ratings* denotes the participants' ratings shifted 200 ms later, a predictor named *L2pitchIC* denotes Pitch IC shifted by 400 ms, and so on. Only optimal models are presented below, selected based on Bayesian Information Criteria (BIC), confidence intervals on fixed-effect predictors, log likelihood ratio tests between pairs of models, correlation tests between models and the data, and the proportion of data squares fit. All analyses were performed using RStudio 1.0.136, running R 3.1.2.

Results

Melody-level Analysis

In this section, we describe analyses of the mean ratings melody by melody; the experiment manipulated the pitch expectancy of the original melodies, to provide a causal test of its influence. Time is treated as a discontinuous variable. Figure 3 shows mean ratings for each melody averaged over time, and important comparisons are highlighted in Figure 4.

Expectancy ratings. There were significant effects of musical training, in which musicians rated melody unexpectedness higher (musicians $M = 4.40$; nonmusicians $M = 4.16$; $F(1, 8343) = 73.12$, $p < .0001$); stimulus modification, in which modified melodies, regardless of direction of manipulation (predictable to unpredictable or vice versa), were rated as more unexpected (original melodies $M = 3.92$; modified melodies $M = 4.65$; $F(1, 8342) = 569.75$, $p < .0001$); and stimulus predictability, in which more predictable melodies were rated with lower unexpectedness than unpredictable melodies (PP melodies $M = 3.48$; PU melodies $M = 4.71$; UP melodies $M = 3.92$; UU melodies $M = 4.66$; $F(3, 8340) = 251.58$, $p < .0001$) on mean expectancy ratings. Pitch predictability and onset predictability were both significant predictors, in which mean ratings for melodies with predictable pitch, unpredictable pitch, predictable onset, and unpredictable onset were 4.09, 4.29, 3.70, and 4.68, respectively ($F(1, 8342) = 83.05$, $p < .0001$ and $F(1, 8342) = 644.31$, $p < .0001$), and the interaction between the two predictors was also significant; there is a more pronounced effect of onset predictability on ratings, $t(8,340) = -7.36$, $p < .0001$. We also investigated the effect of stimulus predictability on ratings for original and modified melodies separately; the means for PP, PU, UP, and UU melodies were 1.88, 4.47, 3.58, and 5.19, respectively, $F(3, 4223) = 1,866.2$, $p < .0001$, and for aPP, aPU, aUP, and aUU melodies were 4.27, 4.16, 5.29, and 4.96, respectively, $F(3, 4112) = 264.36$, $p < .0001$. The two global models confirmed nearly all the aforementioned results, producing two additional findings: Pitch, $t(8,336) = -3.76$, $p = .0001$, and IOI, $t(8,336) = -3.72$, $p = .0001$, were significant predictors in both global models, and pitch predictability became insignificant in its model, $t(2) = 0.24$, $p = .80$. In summary, all predictors of interest were significant, including the interaction between pitch predictability and onset predictability. These results are largely replicated using noninterpolated data, in which only pitch is no longer a significant predictor, $t(2,218) = 1.05$, $p = .29$.

Arousal ratings. There were significant effects of musical training in which musicians rate melodies as more arousing overall as compared with nonmusicians (musicians $M = 118.16$; nonmusicians $M = 112.90$; $F(1, 8017) = 25.30$, $p < .0001$), and stimulus predictability in which more predictable melodies were rated as more arousing (PP melodies $M = 151.73$; PU melodies $M = 109.45$; UP melodies $M = 128.86$; UU melodies $M = 95.95$; $F(3, 8015) = 667.31$, $p < .0001$). There was no effect of stimulus modification in either direction of manipulation (original melodies $M = 115.83$; modified melodies $M = 115.27$; $F(1, 8017) = .62$, $p = .42$). Pitch predictability and onset predictability were both significant predictors in which mean ratings for melodies with predictable pitch, unpredictable pitch, predictable onset, and unpredictable onset were 125.29, 112.40, 135.29, and 102.7, respec-

tively, $F(1, 8017) = 208.38$, $p < .0001$ and $F(1, 8017) = 1,804.3$, $p < .0001$, and the interaction between the two predictors was not significant here, though similarly to expectancy ratings, onset predictability still has a larger effect on mean ratings than pitch predictability, $t(8,015) = 1.08$, $p = .28$. Stimulus predictability was also a significant predictor when original and artificial melodies' ratings were investigated separately, with ratings for PP, PU, UP, and UU melodies averaging 138.62, 111.14, 121.07, and 100.79, respectively, $F(3, 3956) = 210.16$, $p < .0001$, and aPP, aPU, aUP, and aUU melodies averaging 137.10, 91.56 144.96, and 107.83, respectively, $F(3, 4054) = 556.76$, $p < .0001$. The two global models confirm all the above results and add pitch, $t(8,011) = -17.72$, $p < .0001$ and IOI, $t(8,011) = 18.58$, $p < .0001$ as significant predictors. In summary, stimulus modification is the only predictor of interest that did not have a significant effect on arousal ratings. These results are largely replicated using noninterpolated data, in which only pitch is no longer a significant predictor, $t(2,186) = -0.99$, $p = .32$.

Valence ratings. There were significant effects of musical training in which musicians overall rated melodies as having lower valence (musicians $M = 81.26$; nonmusicians $M = 84.08$; $F[1, 8017] = 5.38$, $p = .02$); stimulus modification, regardless of direction of manipulation, in which original melodies had more positive valence than artificial melodies (original melodies $M = 91.20$; artificial melodies $M = 74.33$; $F[1, 8017] = 206.84$, $p < .0001$); and stimulus predictability in which more predictable melodies are rated more positively than unpredictable melodies (PP melodies $M = 109.87$; PU melodies $M = 74.00$; UP melodies $M = 87.00$; UU melodies $M = 70.02$; $F(3, 8015) = 224.81$, $p < .0001$). Pitch predictability and onset predictability were both significant predictors in which mean ratings for melodies with predictable pitch, unpredictable pitch, predictable onset, and unpredictable onset were 91.93, 78.51, 98.43, and 72.01, respectively, $F(1, 8017) = 122.51$, $p < .0001$ and $F(1, 8017) = 559.04$, $p < .0001$, and the interaction between the two predictors was significant, in which onset predictability again has a larger effect on mean ratings than pitch predictability, $t(8,015) = 8.40$, $p < .0001$. Stimulus predictability is also a significant predictor when investigating original and artificial melodies separately, in which PP, PU, UP, and UU melodies have mean arousal ratings of 171.90, 77.96, 94.59, and 44.46, respectively, $F(3, 3956)$, 1,582.6, $p < .0001$, and aPP, aPU, aUP, and aUU melodies have mean ratings of 78.98, 93.21, 45.66, and 70.19, respectively, $F(3, 4054) = 276.84$, $p < .0001$. The two global models include IOI, $t(8,011) = 22.07$, $p < .0001$, but not pitch, $t(8,011) = -1.48$, $p = .13$, as significant predictors (in both models) and remove pitch predictability, $t(8,011) = 0.90$, $p = .36$, from the set of significant predictors found above. In summary, all predictors of interest are significant, including the interaction between pitch predictability and onset predictability. These results are largely replicated using noninterpolated data, in which only pitch and musical training are no longer significant predictors, $t(2,186) = -0.70$, $p = .48$, and $t(2,186) = 0.69$, $p = .48$, respectively.

This melody-level analysis has demonstrated that musical training and stimulus predictability predict expectancy, arousal, and valence ratings, with only one exception, in which musical training does not predict valence ratings when these are averaged for each event. Furthermore, there is a significant interaction between pitch predictability and onset predictability for expectancy and valence
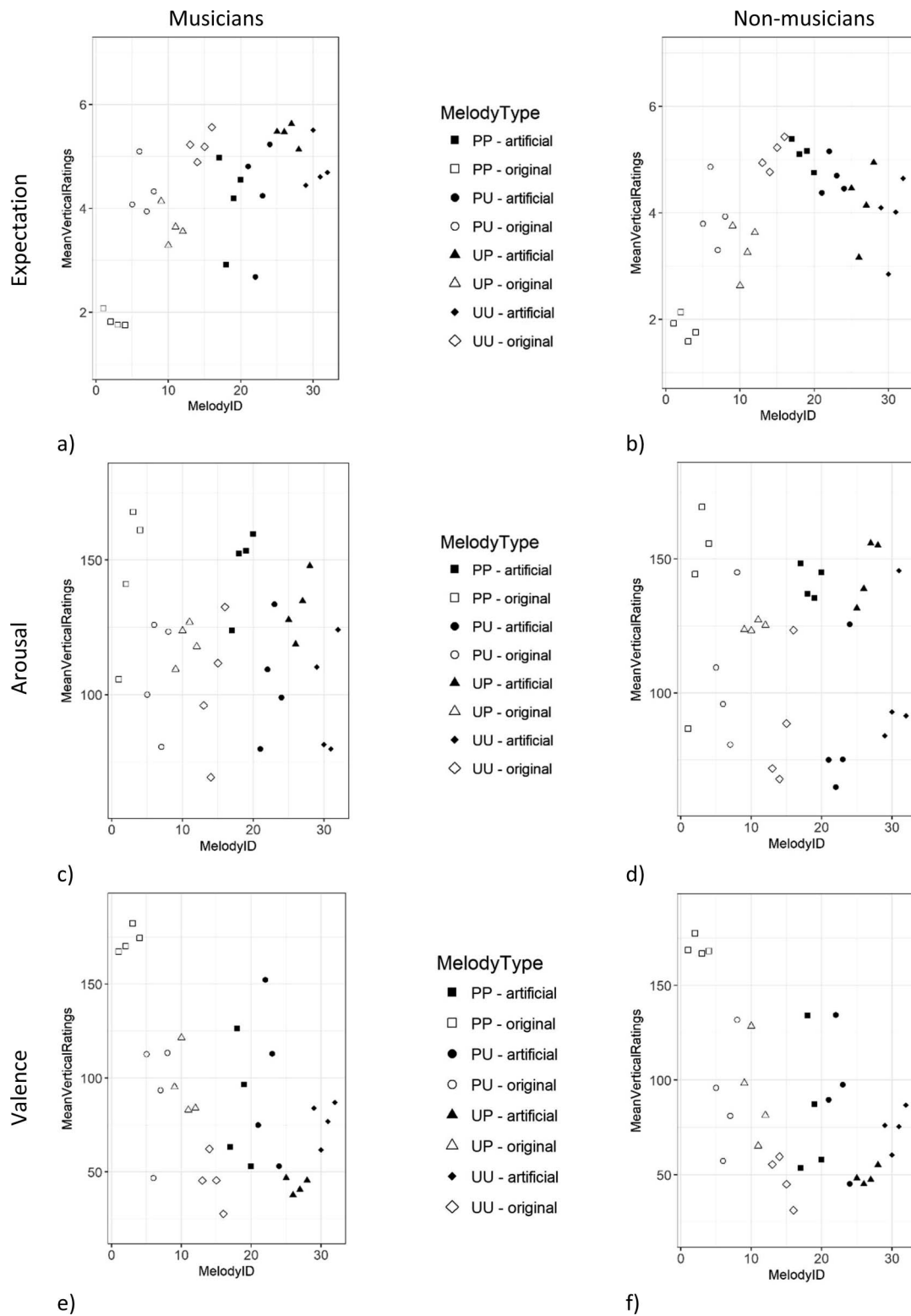
Figure 3. Mean expectancy (a and b), arousal (c and d), and valence (e and f) ratings for each melody for musicians (a, c, e) and nonmusicians (b, d, f). Hollow shapes illustrate original melodies, whereas filled shapes illustrate artificial melodies.

ratings, and a similar pattern for arousal ratings, in which onset predictability has a larger effect on ratings than pitch predictability. Stimulus modification is a significant predictor for expectancy and valence ratings only. In the next section, the results of a CSTSA will be presented.

Cross-sectional Time Series Analysis

Here we present the analyses of the continuous time-series data resulting from participants' ongoing responses throughout listening to the melodies.

Expectancy, arousal, and valence ratings were modeled separately using mixed-effects autoregressive models with random intercepts on Participant ID and Melody ID, as well as random slopes on the fixed effect predictor with the largest coefficient before slopes were added. Fixed-effect predictors were *time*, *musical training*, *stimulus predictability*, *stimulus modification*, autoregressive lags of up to 15 (in which each lag represents 200 ms, for a total of 3 s), and exogenous lags of pitch and onset IC of up to 15. A combined pitch and onset information predictor was also tested to evaluate whether a combined measure superseded the separate pitch and onset IC predictors. Maximum lags for consideration were selected based on a previously reported rate of change of emotional responses (Juslin & Västfjäll, 2008) and precedent in this type of analysis (Dean et al., 2014a). Pitch and IOI were
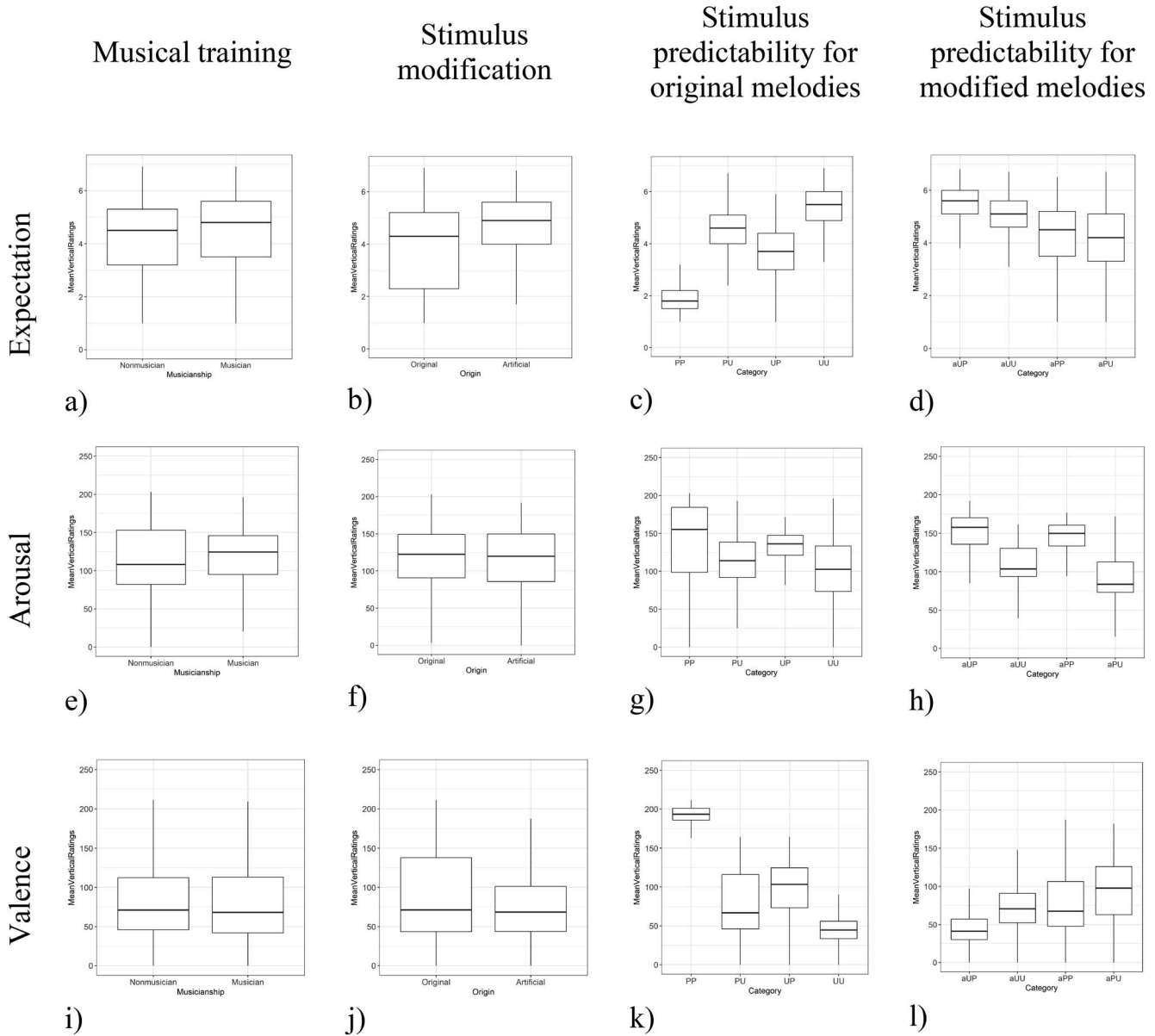
Figure 4. Box plots illustrating important mean comparisons between musicians and nonmusicians (a, e, i), original and modified melodies (b, f, j), stimulus predictability categories for original (c, g, k), and modified (d, h, l) melodies for expectation (a–d), arousal (e–h), and valence (i–l) ratings.
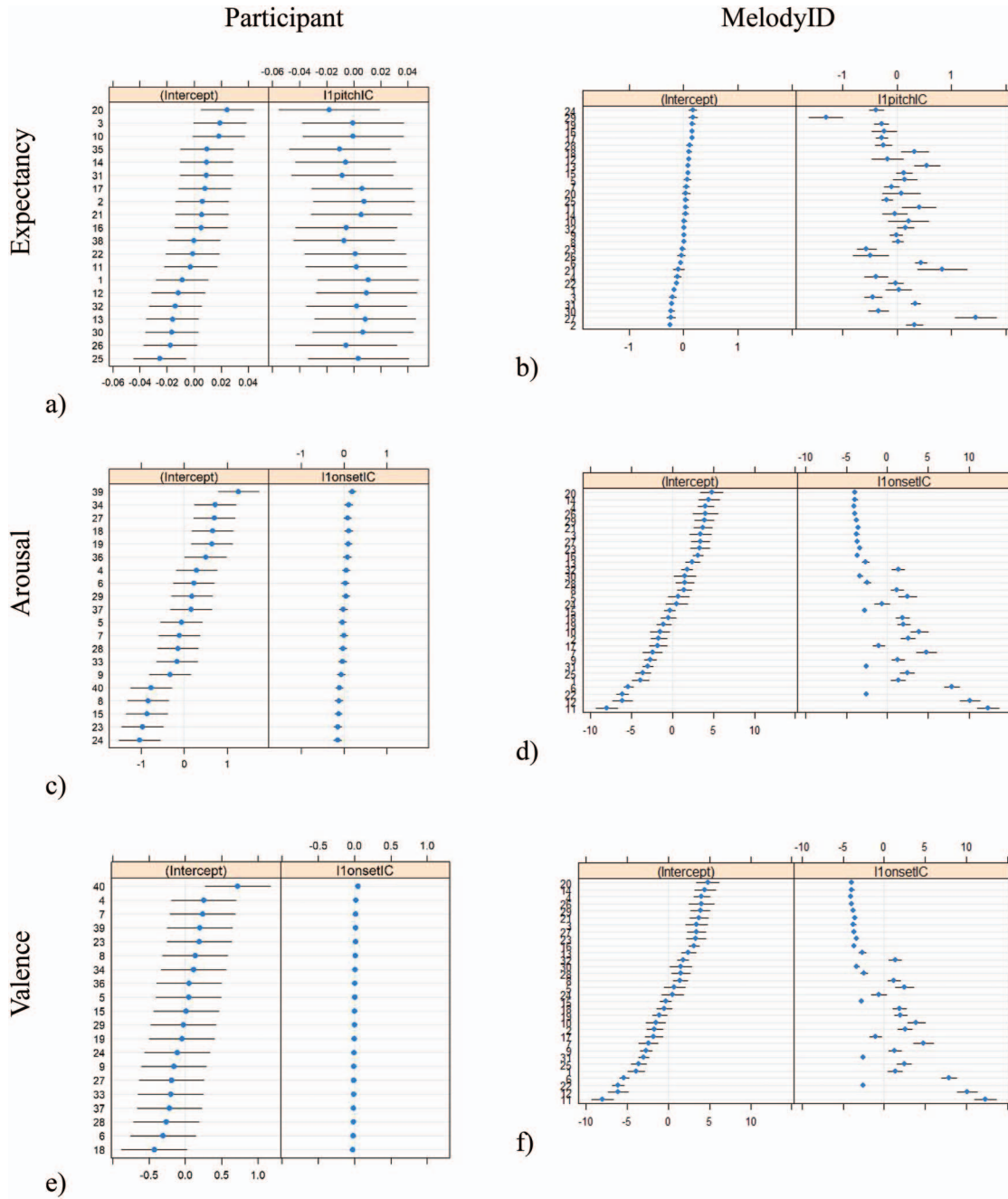
Figure 5. Intercept (left) and slope (right) values of random effects on Participant (a, c, e) and Melody identity (ID; b, d, f) for expectancy (a, b), arousal (c, d), and valence (e, f) models. These show how each individual participant and melody was modeled and illustrate the variance among participants and melodies. See the online article for the color version of this figure.

subsequently added as fixed-effect predictors to investigate the potential effects of musical structure on ratings (which might be in part through an emotional contagion mechanism). See Figures 5 and 6 for an illustration of the variance fitted by random effects and the fit of the models for a selection of melodies and participants.

Expectancy ratings. The best CSTSA model for expectancy ratings is summarized in Table B1 in Appendix B. In this model,

although autoregression and random effects were duly considered, an effect of musicianship was still clearly observed in addition to those of pitch IC and onset IC and the optimal selection of their lags. Thus, the model included random intercepts and random slopes for L1pitchIC on Melody ID and Participant ID, as well as fixed effects of musicianship, L = 0–1, 7–8 of pitch IC; L = 0–2, 10, 12 of onset IC; and p = 1–2, 4–6, 15 of autoregression. All predictors were significant, as Wald 95% confidence intervals did
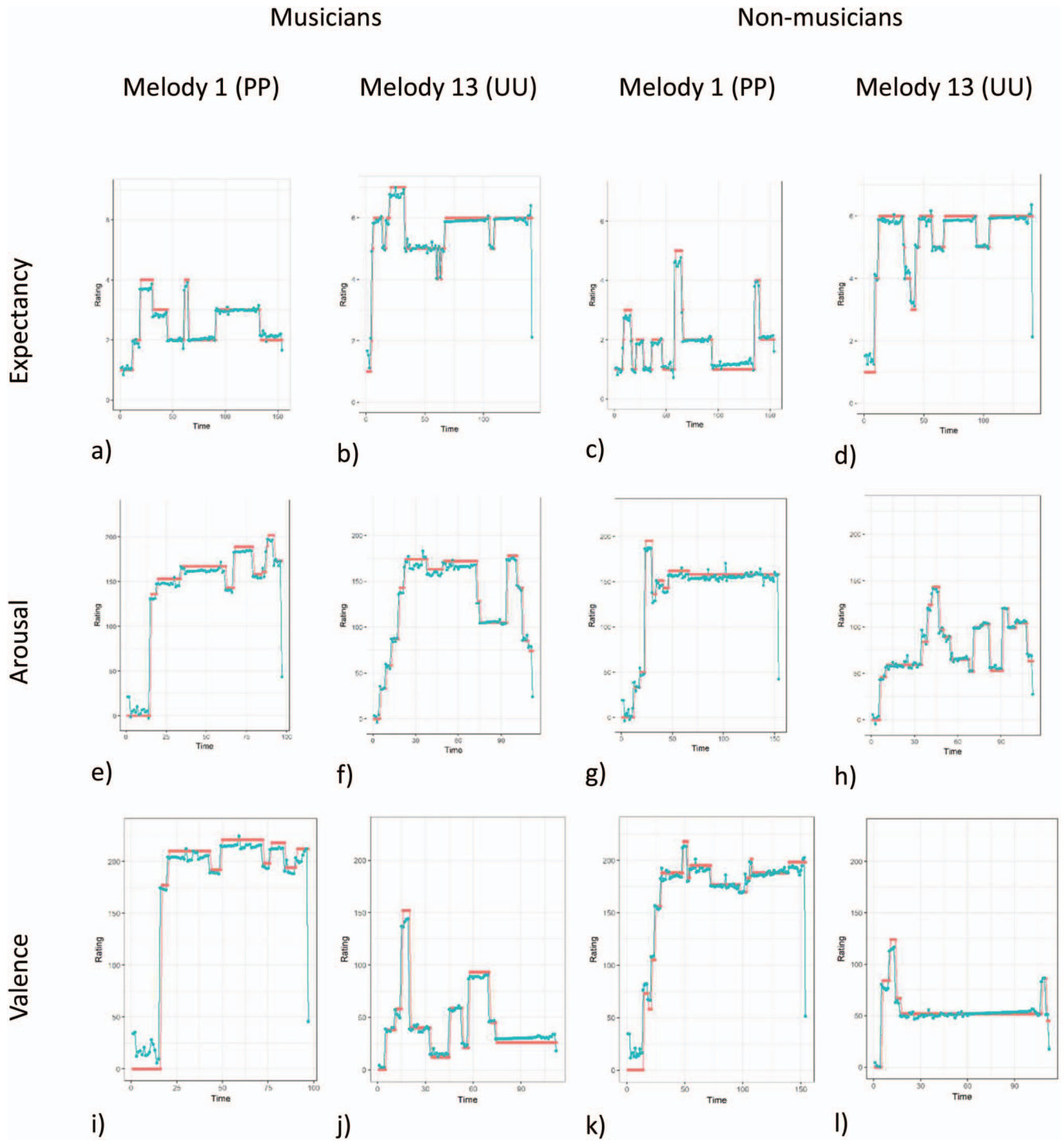
Figure 6. Expectancy (a–d), arousal (e–h), and valence (i–l) ratings for single randomly selected participants (six musicians [a, b, e, f, i, j; Participants 14, 35, 34, 18, 27, and 7] and six nonmusicians [c, d, g, h, k, l; Participants 1, 10, 8, 33, 5, and 37]) are plotted for Melodies 1 (a, c, e, g, i, k) and 13 (b, d, f, h, j, l), examples of PP and UU categories, respectively. Ratings predicted by the model (teal) for those melodies for each of those participants only (single extracts) are plotted alongside their actual ratings (pink). Residuals were too small to illustrate on the same plot. These plots illustrate the high explanatory power of our model due to its random effects structure fitted specifically to this data set. See the online article for the color version of this figure.

not include zero. The addition of stimulus predictability as a fixed effect did not improve the model, $\chi^2(3) = 1.80$, $p = .61$, whereas musicianship and stimulus modification did, $\chi^2(2) = 13.36$, $p = .001$ and $\chi^2(1) = 3.91$, $p = .04$, respectively. The further addition of pitch and IOI significantly improved the model, $\chi^2(2) = 409.33$, $p < .0001$, and removed stimulus modification as a significant predictor. Combined pitch and onset information content with lags of pitch and onset from the best model outlined in the abovementioned text was significantly worse, $\chi^2(6) = 972.6$, $p < .0001$.

A correlation test between the data and the model is highly significant, with correlation .93, $t(82,486) = 783.09$, $p < .0001$. A proportion of data squares fit test is also high, with the model explaining 98% of the data. Although this particular model did not converge, a model without random slopes removed did converge where all fixed effects were significant, model fit was equally good, correlation test: .93, $t(82,486) = 780.53$, $p < .0001$; proportion of data squares fit: 98%, and the inclusion of slopes improved the model significantly; therefore random slopes were reinserted into the best model as per the experimental design (Barr et al., 2013). The final model thus includes design-driven random effects, musicianship, stimulus modification, pitch, IOI, optimal autoregressive lags of expectancy ratings, and optimal lags of pitch IC and onset IC. These results are replicated using noninterpolated data, with only the selection of lags differing.

Arousal ratings. The best CSTSA model for arousal ratings is summarized in Table B2 in Appendix B. This model revealed stimulus predictability as a significant predictor of arousal ratings in addition to pitch IC and onset IC, and a selection of their lags when autoregression and random effects were considered. The model included random intercepts and random slopes for L1onsetIC on Melody ID and Participant ID, as well as fixed effects L = 0–1, 6–8, 10–13, 15 of pitch IC; L = 0–4, 7, 10, 12–15 of onset IC; and $p = 1, 3, 5$–6, 15 of autoregression. All predictors were significant, as Wald 95% confidence intervals did not include zero. The addition of musicianship and stimulus modification as fixed effects did not improve the model, $\chi^2(2) = 0.60$, $p = .74$, and $\chi^2(2) = 1.72$, $p = .42$, respectively, whereas stimulus predictability did, $\chi^2(2) = 14.91$, $p = .0005$. The further addition of pitch and IOI significantly improved the model, $\chi^2(2) = 178.89$, $p < .0001$, in which both are significant predictors of arousal ratings. Combined pitch and onset information content with lags of pitch and onset from the best model outlined in the abovementioned text was significantly worse, $\chi^2(13) = 4,482.2$, $p < .0001$.

A correlation test between the data and the model is highly significant, with correlation .96, $t(80,183) = 978.48$, $p < .0001$. A proportion of data squares fit test is also high, with our model explaining 98% of the data. Although this particular model did not converge, a model without random slopes removed did converge where all fixed effects were significant, model fit was equally good, correlation test: .95, $t(80,183) = 959.73$, $p < .0001$; proportion of data squares fit: 98%, and the inclusion of slopes improved the model significantly, $\chi^2(5) = 335.3$, $p < .0001$; therefore random slopes were reinserted into the best model as per the experimental design (Barr et al., 2013). The final model thus includes design-driven random effects, stimulus predictability, pitch, IOI, optimal autoregressive lags of expectancy ratings, and optimal lags of pitch IC and onset IC. These results are replicated

using noninterpolated data, with only the selection of lags differing.

Valence ratings. The best CSTSA model for valence ratings is summarized in Table B3 in Appendix B. This model revealed significant effects of only pitch IC and onset IC, and a selection of their lags when autoregression and random effects were considered. The model included random intercepts and random slopes for L1onsetIC on Melody ID and Participant ID, as well as fixed effects L = 0–1, 5, 8–9, 11–13, 15 of pitch IC; L = 0–1, 3–4, 10, 13 of onset IC; and $p = 0, 3$–7, 9, 15 of autoregression. All predictors were significant, as Wald 95% confidence intervals did not include zero. The addition of musicianship, stimulus predictability, and modification as fixed effects did not improve the model, $\chi^2(1) = 0.29$, $p = .58$, $\chi^2(3) = 4.77$, $p = .18$, and $\chi^2(1) = 3.46$, $p = .06$, respectively. The further addition of pitch and IOI significantly improved the model, $\chi^2(1) = 600.99$, $p < .0001$, in which both are significant predictors of arousal ratings. Combined pitch and onset information content with lags of pitch and onset from the best model outlined in the abovementioned text was significantly worse, $\chi^2(10) = 194.72$, $p < .0001$.

A correlation test between the data and the model is highly significant, with correlation .94, $t(80,183) = 827.83$, $p < .0001$. A proportion of data squares fit test is also high, with our model explaining 98% of the data. Although this particular model did not converge, a model without random slopes removed did converge where all fixed effects were significant, model fit was equally good, correlation test: .94, $t(80,183) = 959.73$, $p < .0001$; proportion of data squares fit: 95%, and the inclusion of slopes improved the model significantly, $\chi^2(4) = 805.25$, $p < .0001$; therefore random slopes were reinserted into the best model as per the experimental design (Barr et al., 2013). The final model thus includes design-driven random effects, pitch, IOI, optimal autoregressive lags of expectancy ratings, and optimal lags of pitch IC and onset IC. These results are replicated using noninterpolated data, with only the selection of lags differing.

Discussion

The results provide answers to all three of our research questions. First, we find evidence that predictability of both pitch and temporal musical structure have an effect on listeners' expectancies and emotional reactions and that these can be manipulated. Second, we find that contrary to a prediction based on complexity, temporal expectancy influences perception more strongly than pitch expectancy. Finally, we find that individual differences generally supersede effects of musical training (Dean et al., 2014a) and intermelody differences were more substantial than differences between melody predictability groups (PP, UP, PU, and UU) or manipulation type, in which differences between predictability groups could nevertheless be detected in the discontinuous, melody-level analysis.

Using IDyOM (Pearce, 2005) to calculate average pitch and onset information content, we classified folk songs into four categories based on overall expectedness, in which average pitch expectancy and average onset expectancy could be high or low. We also manipulated pitch expectancy to transform expected pitches into unexpected ones and vice versa. The four melody categories resulted in different subjective ratings of expectancy, arousal, and valence; high pitch and onset information content

arousing than PU and UU (low onset expectancy), matches this interpretation, which is further supported by previous research establishing a positive correlation between tempo and arousal (Carpentier & Potter, 2007; Husain, Thompson, & Schellenberg, 2002). The significant effect of pitch on ratings is more surprising; a pattern of higher average pitch for PP and UP categories corresponds to lower unexpectedness ratings, higher arousal ratings, and higher valence ratings for these categories as compared with PU and UU categories. However, coefficients for pitch and IOI are smaller than almost all other predictors in expectancy, arousal, and valence models, suggesting that their overall influence is minimal compared with pitch and onset IC on subjective expectancy and emotion responses.

Also, similar to Dean et al. (2014a), the use of CSTSA allows us to evaluate evidence for the presence of a common perceptual mechanism across all pieces of music heard. To do this, predictors encoding melodies by stimulus predictability and modification were added to the basic models, in which a null effect of these additional predictors would indicate that the type of melody does not matter and the listeners' ratings depend only on pitch and onset IC in all melodies. In the case of valence ratings, neither stimulus predictability nor stimulus modification were found to provide any additional predictive power to the model, whereas stimulus modification was a helpful predictor for expectancy ratings and stimulus predictability for arousal ratings. However, explanations were proposed for these results, and our data provide some support for a common perceptual mechanism across all melodies.

Relative Salience

Having considered the relative importance of pitch and onset IC in the context of our models of participant expectancy, arousal, and valence ratings, here we consider how this relates to relative perceptual salience. The question of relative perceptual salience between musical parameters such as pitch, timing, structure, and harmony in music cognition is important but challenging and lacks a unified explanation (Dibben, 1999; Esber & Haselgrove, 2011; Prince et al., 2009; Uhlig, Fairhurst, & Keller, 2013). Generally, pitch or melody is considered the most salient aspect of a piece of music. Prince et al. (2009), for example, argued that there are many more possible pitches than there are rhythmic durations or chords; therefore, pitch takes more attentional resources to process and is more salient. On the other hand, in a corpus analysis of 18th- and 19th-century string quartets, Duane (2013) found that onset and offset synchrony were the most important predictors of streaming perception of these string quartets, with pitch explaining half the variance that onset and offset synchrony did and harmonic overlap explaining an almost insignificant amount. It is also important to consider the musical genre when discussing salience, as certain genres are more rhythmically driven (i.e., rap, electronic dance music, and African drum music), whereas others are more melodically driven (i.e., opera). Folk music is more ambivalent and may vary song by song. Other genres may well produce different results; something that would be worth exploring in the future. Our stimuli best fit Prince et al.'s (2009) description of musical salience, as these melodies contain more different pitches than different rhythmic values. This would imply that the pitch dimension is more complex and therefore more salient. However, our results indicate that onset IC is more salient than pitch IC,

though here we evaluate the perception of emotion alongside the subjective experience of expectancy as opposed to auditory streaming (Duane, 2013; Prince et al., 2009). Interestingly, work in cue salience in the context of associative learning explores the effect of predictability and uncertainty on salience (Esber & Haselgrove, 2011), with one model predicting increased salience for cues with high predictability (Mackintosh, 1975) and another model predicting increased salience for cues with high uncertainty (Pearce & Hall, 1980). Though contradictory, these models have each accumulated significant evidence and have more recently led to the development of both hybrid (Pearce & Mackintosh, 2010) and new unified models of cue salience (Esber & Haselgrove, 2011). We considered the possibility that high and low uncertainty and pitch and onset lag coefficients interacted, so that melodies with high-pitch predictability (expectancy) and low-onset predictability (PU) led to larger pitch IC coefficients than onset IC coefficients, and vice versa. This effect was not found in the data (see Appendix C), so we conclude that in this particular paradigm, onset is the more salient cue overall.

A Mechanism for Emotional Induction

Returning to the identified lack of research into specific mechanisms for emotional induction by music (Juslin & Västfjäll, 2008; Meyer, 1956), the present research makes a single but significant step toward isolating individual mechanisms. The study explicitly controlled for four of the eight proposed mechanisms, considered two unlikely, and manipulated one while considering another as a covariate. Brain stem reflexes, evaluative conditioning, episodic memory, and visual imagery are controlled for by presenting novel stimuli with equal tempo, intensity, and timbre alongside a rating task. Rhythmic entrainment and cognitive appraisal are highly unlikely due to the lack of driving rhythm and experimental setting. Emotional contagion, information conveyed by musical structure itself, was addressed by including pitch and duration values into our CSTSA models of expectancy, arousal, and valence ratings. Though these were significant predictors, they carried less weight than the lags of IC predictors. We examined musical expectancy by selecting stimuli with either high or low pitch and onset expectancy and additionally explicitly manipulated pitch expectancy, finding evidence for a consistent effect of pitch and onset expectancy on ratings of arousal and valence by musicians and nonmusicians. We additionally find that onset is more salient than pitch and that musicians give higher unexpected ratings than nonmusicians, but group differences are overridden by individual differences on emotion ratings. Potential future work includes the use of stimuli at less extreme ends of the expectancy spectrum to validate these findings and produce more generalizable models, manipulating onset IC in addition to pitch IC, allowing the evaluation of dependencies between the two (Palmer & Krumhansl, 1987b), exploring interactions of predictability and entropy on salience cues in emotion ratings, and investigating other potential emotional induction mechanisms in a similarly controlled way, working toward an integrated model of musical emotion induction and perception.

References

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal*

of Memory and Language, 68, 255–278. http://dx.doi.org/10.1016/j.jml .2012.11.001

Bauer, A.-K. R., Jaeger, M., Thorne, J. D., Bendixen, A., & Debener, S. (2015). The auditory dynamic attending theory revisited: A closer look at the pitch comparison task. *Brain Research, 1626,* 198–210. http://dx .doi.org/10.1016/j.brainres.2015.04.032

Bittman, B., Croft, D. T., Jr., Brinker, J., van Laar, R., Vernalis, M. N., & Ellsworth, D. L. (2013). Recreational music-making alters gene expression pathways in patients with coronary heart disease. *Medical Science Monitor, 19,* 139–147. http://dx.doi.org/10.12659/MSM.883807

Boltz, M. G. (1999). The processing of melodic and temporal information: Independent or unified dimensions? *Journal of New Music Research, 28,* 67–79. http://dx.doi.org/10.1076/jnmr.28.1.67.3121

Brattico, E., Näätänen, R., & Tervaniemi, M. (2001). Context effects on pitch perception in musicians and nonmusicians: Evidence from event-related-potential recordings. *Music Perception, 19,* 199–222. http://dx .doi.org/10.1525/mp.2001.19.2.199

Carey, D., Rosen, S., Krishnan, S., Pearce, M. T., Shepherd, A., Aydelott, J., & Dick, F. (2015). Generality and specificity in the effects of musical expertise on perception and cognition. *Cognition, 137,* 81–105. http:// dx.doi.org/10.1016/j.cognition.2014.12.005

Carlsen, J. C. (1981). Some factors which influence melodic expectancy. *Psychomusicology: A Journal of Research in Music Cognition, 1,* 12–29. http://dx.doi.org/10.1037/h0094276

Dillman Carpenter, F. R., & Potter, R. F. (2007). Effects of music on physiological arousal: Explorations into tempo and genre. *Media Psychology, 10,* 339–363. http://dx.doi.org/10.1080/15213260701533045

Castillo-Pérez, S., Gómez-Pérez, V., Velasco, M. C., Pérez-Campos, E., & Mayoral, M.-A. (2010). Effects of music therapy on depression compared with psychotherapy. *The Arts in Psychotherapy, 37,* 387–390. http://dx.doi.org/10.1016/j.aip.2010.07.001

Cirelli, L. K., Spinelli, C., Nozaradan, S., & Trainor, L. J. (2016). Measuring neural entrainment to beat and meter in infants: Effects of music background. *Frontiers in Neuroscience, 10,* 229. http://dx.doi.org/10 .3389/fnins.2016.00229

Collins, T., Laney, R., Willis, A., & Garthwaite, P. H. (2011). Modeling pattern importance in Chopin's Mazurkas. *Music Perception, 28,* 387–414. http://dx.doi.org/10.1525/mp.2011.28.4.387

Dean, R. T., Bailes, F., & Dunsmuir, W. T. M. (2014a). Shared and distinct mechanisms of individual and expertise-group perception of expressed arousal in four works. *Journal of Mathematics and Music, 8,* 207–223. http://dx.doi.org/10.1080/17459737.2014.928753

Dean, R. T., Bailes, F., & Dunsmuir, W. T. M. (2014b). 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, 8,* 183–205. http://dx.doi.org/10 .1080/17459737.2014.928752

Dibben, N. (1999). The perception of structural stability in atonal music: The influence of salience, stability, horizontal motion, pitch commonality, and dissonance. *Music Perception, 16,* 265–294. http://dx.doi.org/ 10.2307/40285794

Duane, B. (2013). Auditory streaming cues in eighteenth- and early nineteenth-century string quartets: A corpus-based study. *Music Perception: An Interdisciplinary Journal, 31,* 46–58. http://dx.doi.org/10.1525/ mp.2013.31.1.46

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 and Behavioral Neuroscience, 13,* 533–553. http://dx.doi.org/10.3758/ s13415-013-0161-y

Esber, G. R., & Haselgrove, M. (2011). Reconciling the influence of predictiveness and uncertainty on stimulus salience: A model of attention in associative learning. *Proceedings of the Royal Society of London B: Biological Sciences, 278,* 2553–2561. http://dx.doi.org/10.1098/rspb .2011.0836

Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., & Pantev, C. (2004). Musical training enhances automatic encoding of melodic contour and interval structure. *Journal of Cognitive Neuroscience, 16,* 1010–1021. http://dx.doi.org/10.1162/0898929041502706

Gingras, B., Pearce, M. T., Goodchild, M., Dean, R. T., Wiggins, G., & McAdams, S. (2016). Linking melodic expectation to expressive performance timing and perceived musical tension. *Journal of Experimental Psychology: Human Perception and Performance, 42,* 594–609. http:// dx.doi.org/10.1037/xhp0000141

Granot, R., & Donchin, E. (2002). Do Re Mi Fa Sol La Ti—Constraints, congruity, and musical training: An event-related brain potentials study of musical expectancies. *Music Perception, 19,* 487–528. http://dx.doi .org/10.1525/mp.2002.19.4.487

Habibi, A., & Damasio, A. (2014). Music, feelings, and the human brain. *Psychomusicology: Music, Mind, and Brain, 24,* 92–102. http://dx.doi .org/10.1037/pmu0000033

Hannon, E. E., Soley, G., & Ullal, S. (2012). Familiarity overrides complexity in rhythm perception: A cross-cultural comparison of American and Turkish listeners. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 543–548. http://dx.doi.org/10.1037/ a0027225

Hannon, E. E., & Trehub, S. E. (2005a). Metrical categories in infancy and adulthood. *Psychological Science, 16,* 48–55. http://dx.doi.org/10.1111/ j.0956-7976.2005.00779.x

Hannon, E. E., & Trehub, S. E. (2005b). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences of the United States of America, 102,* 12639–12643. http://dx.doi.org/10.1073/pnas.0504254102

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

Huron, D. (2001). What is a musical feature? *Music Theory Online, 7,* 1–18.

Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation.* Cambridge, MA: The MIT Press.

Husain, G., Thompson, W. F., & Schellenberg, E. G. (2002). Effects of musical tempo and mode on arousal, mood, and spatial abilities. *Music Perception, 20,* 151–171. http://dx.doi.org/10.1525/mp.2002.20.2.151

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review, 96,* 459–491. http://dx.doi.org/10.1037/ 0033-295X.96.3.459

Jones, M. R., Boltz, M., & Kidd, G. (1982). Controlled attending as a function of melodic and temporal context. *Perception and Psychophysics, 32,* 211–218. http://dx.doi.org/10.3758/BF03206225

Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science, 13,* 313–319. http://dx.doi.org/10.1111/1467-9280 .00458

Juslin, P. N., Liljeström, S., Västfjäll, D., & Lundqvist, L.-O. (2011). How does music evoke emotions? Exploring the underlying mechanisms. In P. N. Juslin & J. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 605–644). New York, NY: Oxford University Press. Retrieved from https://philpapers.org/rec/JUSHDM

Juslin, P. N., & Sloboda, J. (2011). *Handbook of music and emotion: Theory, research, applications.* New York, NY: Oxford University Press. Retrieved from https://books.google.co.uk/books?hl=en&lr= &id=e250IFNvw1sC&oi=fnd&pg=PT6&dq=music+and+emotion& ots=U_cVJdhueJ&sig=WJjyrykHIdsA_xfMrv11w0GKDrE

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

Khalfa, S., Bella, S. D., Roy, M., Peretz, I., & Lupien, S. J. (2003). Effects of relaxing music on salivary cortisol level after psychological stress. *Annals of the New York Academy of Sciences, 999,* 374–376. http://dx .doi.org/10.1196/annals.1284.045

Khalfa, S., Isabelle, P., Jean-Pierre, B., & Manon, R. (2002). Event-related skin conductance responses to musical emotions in humans. *Neuroscience Letters, 328,* 145–149.

Koelsch, S., Fritz, T., & Schlaug, G. (2008). Amygdala activity can be modulated by unexpected chord functions during music listening. *NeuroReport: For Rapid Communication of Neuroscience Research, 19,* 1815–1819. http://dx.doi.org/10.1097/WNR.0b013e32831a8722

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. (2001). *Cognitive foundations of musical pitch.* New York, NY: Oxford University Press.

Lerdahl, F. (1989). Atonal prolongational structure. *Contemporary Music Review, 4,* 65–87. http://dx.doi.org/10.1080/07494468900640211

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music.* Cambridge, Massachusetts: MIT Press.

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review, 82,* 276–298. http://dx.doi.org/10.1037/h0076778

Meyer, L. (1956). *Emotion and meaning in music.* Chicago, IL: University of Chicago Press. Retrieved from http://www.press.uchicago.edu/ucp/books/book/chicago/E/bo3643659.html

Narmour, E. (1991). The top-down and bottom-up systems of musical implication: Building on Meyer's theory of emotional syntax. *Music Perception, 9,* 1–26. http://dx.doi.org/10.2307/40286156

Palmer, C., & Krumhansl, C. L. (1987a). Independent temporal and pitch structures in determination of musical phrases. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 116–126. http://dx.doi.org/10.1037/0096-1523.13.1.116

Palmer, C., & Krumhansl, C. L. (1987b). Pitch and temporal contributions to musical phrase perception: Effects of harmony, performance timing, and familiarity. *Perception and Psychophysics, 41,* 505–518. http://dx .doi.org/10.3758/BF03210485

Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 728–741.

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review, 87,* 532–552. http://dx.doi.org/10.1037/0033-295X.87.6.532

Pearce, J. M., & Mackintosh, N. J. (2010). Two theories of attention: A review and a possible integration. In C. J. Mitchell & M. E. L. Pelley (Eds.), *Attention and associative learning: From brain to behaviour* (pp. 11–39). Oxford, UK: Oxford University Press.

Pearce, M. T. (2005, December). *The construction and evaluation of statistical models of melodic structure in music perception and composition* (Unpublished doctoral dissertation). City University London, London, UK. Retrieved from http://openaccess.city.ac.uk/8459/

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. 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

Pelletier, C. L. (2004). The effect of music on decreasing arousal due to stress: A meta-analysis. *Journal of Music Therapy, 41,* 192–214. http://dx.doi.org/10.1093/jmt/41.3.192

Prince, J. B., Thompson, W. F., & Schmuckler, M. A. (2009). Pitch and time, tonality and meter: How do musical dimensions combine? *Journal of Experimental Psychology: Human Perception and Performance, 35,* 1598–1617. http://dx.doi.org/10.1037/a0016456

Rentfrow, P. J., & Gosling, S. D. (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology, 84,* 1236–1256. http://dx.doi.org/10 .1037/0022-3514.84.6.1236

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110,* 145–172. http://dx.doi.org/10 .1037/0033-295X.110.1.145

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

Schellenberg, E. G. (1996). Expectancy in melody: Tests of the implication-realization model. *Cognition, 58,* 75–125. http://dx.doi.org/ 10.1016/0010-0277(95)00665-6

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

Strait, D. L., Kraus, N., Skoe, E., & Ashley, R. (2009). Musical experience and neural efficiency – effects of training on subcortical processing of vocal expressions of emotion. *European Journal of Neuroscience, 29,* 661–668. http://dx.doi.org/10.1111/j.1460-9568.2009.06617.x

Temperley, D. (1999). What's key for key? The Krumhansl-Schmuckler key-finding algorithm reconsidered. *Music Perception: An Interdisciplinary Journal, 17,* 65–100. http://dx.doi.org/10.2307/40285812

Thompson, W. F., Graham, P., & Russo, F. A. (2005). Seeing music performance: Visual influences on perception and experience. *Semiotica, 156,* 203–227.

Trainor, L. J., & Trehub, S. E. (1992). A comparison of infants' and adults' sensitivity to western musical structure. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 394–402. http://dx .doi.org/10.1037/0096-1523.18.2.394

Uhlig, M., Fairhurst, M. T., & Keller, P. E. (2013). The importance of integration and top-down salience when listening to complex multi-part musical stimuli. *NeuroImage, 77,* 52–61. http://dx.doi.org/10.1016/j .neuroimage.2013.03.051

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., & Levitin, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition, 101,* 80–113. http://dx.doi.org/10.1016/j.cognition .2005.09.003

Appendix A

Summary of Originial Melodies

Table A1
Summary of 16 Original Melodies Used in This Experiment

File name	Data set of origin	Number of events	Average pitch (60 = midC)	Average note duration (24 = quarter)	Mean pitch IC	Mean onset IC	Stimulus predictability
Kindr138	15	33	67.69	74.18	1.3624	0.8962	PP
A162	18	21	70.23	27.42	1.4328	0.8955	PP
Kindr151	15	51	66.05	22.82	1.5971	0.8114	PP
Kindr162	15	19	68.36	26.52	1.5574	0.9254	PP
Deut3480	10	19	72.89	36.94	2.4272	4.4488	PU
Jugos052	4	54	66.22	6.77	2.2543	3.7433	PU
I0511	23	53	66.83	11.67	2.0089	2.4660	PU
Deut3284	10	67	69.71	6.52	2.0913	2.5380	PU
I0533	23	39	67.76	11.79	5.6137	1.1401	UP
A120	18	35	64.05	17.31	5.2750	1.3358	UP
Oestr045	6	30	68.90	36.40	4.7200	1.1290	UP
Oestr046	6	35	64.40	32.22	4.6734	1.1983	UP
Deut3244	10	39	67.64	21.84	3.0216	4.7589	UU
Deut3206	10	52	68.67	22.15	2.9122	4.5098	UU
Deut3437	10	29	71.62	19.86	3.0114	4.3796	UU
Deut3524	10	38	72.76	15.15	2.8472	4.3009	UU

Note. IC = information content; PP = predictable pitch and predictable onset; PU = predictable pitch and unpredictable onset; UP = unpredictable pitch and predictable onset; UU = unpredictable pitch and unpredictable onset.

(Appendices continue)

Appendix B

Results of CSTSA Analysis

Table B1
*Cross-Sectional Time Series Analysis Modelling of Expectancy Ratings for All Melodies;
Coefficients for Fixed Width 95% CIs and Variance of Random Effects*

Model component	Predictor	Coefficient	95% CI
Fixed effects	Intercept	0.307	[0.251, 0.364]
	Time	0.001	[0.001, 0.001]
	Musicianship	0.030	[0.014, 0.047]
	Pitch	−0.002	[−0.002, −0.001]
	IOI	0.003	[0.003, 0.003]
	L1ratings	0.960	[0.953, 0.967]
	L2ratings	−0.065	[−0.073, −0.058]
	L4ratings	−0.061	[−0.069, −0.053]
	L5ratings	0.015	[0.006, 0.025]
	L6ratings	0.035	[0.023, 0.037]
	L15ratings	0.015	[0.012, 0.018]
	PitchIC	−0.263	[−0.309, −0.217]
	L1pitchIC	0.486	[0.306, 0.666]
	L7pitchIC	0.123	[0.079, 0.167]
	L8pitchIC	−0.059	[−0.103, −0.016]
	OnsetIC	−0.731	[−0.794, −0.667]
	L1onsetIC	0.845	[0.769, 0.920]
	L2onsetIC	−0.181	[−0.240, −0.123]
	L10onsetIC	−0.084	[−0.129, −0.039]
	L12onsetIC	0.138	[0.092, 0.183]
	Predictor	Variance	—
Random effects on individuals	Intercept	0.000	—
	L1pitchIC	0.000	—
Random effects on melody ID	Intercept	0.019	—
	L1pitchIC	0.245	—
Residual variance		0.421	—

Note. CI = confidence interval; IC = information content; ID = identity; IOI = interonset interval.

(*Appendices continue*)

Table B2

Cross-Sectional Time Series Analysis Modelling of Arousal Ratings for All Melodies; Coefficients for Fixed Width 95% CIs and Variance of Random Effects

Model component	Predictor	Coefficient	95% CI
Fixed effects	(Intercept)	1.98	[−0.07, 4.03]
	Time	0.06	[0.06, 0.07]
	PU	−3.42	[−5.77, −1.06]
	UP	−0.50	[−2.86, 1.85]
	UU	−4.53	[−6.88, −2.17]
	Pitch	−0.04	[−0.05, −0.03]
	IOI	−0.03	[−0.03, −0.02]
	L1ratings	0.95	[0.94, 0.95]
	L3ratings	0.01	[0.00, 0.01]
	L5ratings	−0.05	[−0.06, −0.05]
	L6ratings	0.03	[0.02, 0.03]
	L15ratings	0.01	[0.01, 0.01]
	PitchIC	−16.6	[−17.7, −15.4]
	L1pitchIC	16.6	[15.5, 17.8]
	L6pitchIC	2.46	[1.31, 3.62]
	L7pitchIC	2.05	[0.70, 3.39]
	L8pitchIC	−2.14	[−3.37, −0.92]
	L10pitchIC	1.86	[0.63, 3.08]
	L11pitchIC	−4.43	[−5.77, −3.10]
	L12pitchIC	4.91	[3.57, 6.25]
	L13pitchIC	−1.95	[−3.18, −0.72]
	L15pitchIC	2.18	[1.23, 3.13]
	OnsetIC	−11.4	[−12.9, −9.83]
	L1onsetIC	72.4	[48.2, 96.6]
	L3onsetIC	6.96	[5.26, 8.66]
	L4onsetIC	−8.38	[−9.98, −6.77]
	L7onsetIC	1.55	[0.34, 2.76]
	L10onsetIC	−6.81	[−8.12, −5.49]
	L12onsetIC	5.43	[3.73, 7.13]
	L13onsetIC	4.47	[2.55, 6.39]
	L14onsetIC	−2.93	[−4.83, −1.04]
	L15onsetIC	3.09	[1.59, 4.58]

	Predictor	Variance	—
Random effects on individuals	Intercept	0.47	—
	L1onsetIC	2.94	—
Random effects on melody ID	Intercept	13.5	—
	L1onsetIC	4,815.2	—
Residual variance		276.7	—

Note. CI = confidence interval; IC = information content; PU = predictable pitch and unpredictable onset; UP = unpredictable pitch and predictable onset; UU = unpredictable pitch and unpredictable onset; ID = identity.

(*Appendices continue*)

Table B3

Cross-Sectional Time Series Analysis Modelling of Valence Ratings for All melodies;
Coefficients for Fixed Width 95% CIs and Variance of Random Effects

Model component	Predictor	Coefficient	95% CI
Fixed effects	(Intercept)	5.38	[3.56, 7.20]
	Time	0.03	[0.03, 0.04]
	Pitch	−0.09	[−0.10, −0.08]
	IOI	0.16	[0.15, 0.18]
	L1ratings	0.92	[0.92, 0.93]
	L3ratings	−0.02	[−0.03, −0.01]
	L4ratings	−0.03	[−0.04, −0.02]
	L5ratings	−0.01	[−0.02, −0.00]
	L6ratings	0.01	[0.00, 0.02]
	L7ratings	0.01	[0.00, 0.02]
	L9ratings	0.00	[0.00, 0.01]
	L15ratings	0.00	[0.00, 0.01]
	PitchIC	−9.19	[−10.6, −7.72]
	L1pitchIC	11.2	[9.74, 12.6]
	L5pitchIC	2.62	[1.45, 3.79]
	L8pitchIC	−3.26	[−4.72, −1.79]
	L9pitchIC	3.29	[1.74, 4.83]
	L11pitchIC	−1.68	[−3.22, −0.15]
	L12pitchIC	2.91	[1.47, 4.83]
	L15pitchIC	1.28	[0.20, 2.36]
	OnsetIC	−20.0	[−22.2, −17.9]
	L1onsetIC	48.5	[29.7, 67.3]
	L3onsetIC	4.05	[1.92, 6.18]
	L4onsetIC	−4.02	[−5.90, −2.13]
	L10onsetIC	−5.35	[−6.65, −4.05]
	L13onsetIC	3.59	[2.32, 4.86]
	Predictor	Variance	—
Random effects on individuals	(Intercept)	0.11	—
	L1onsetIC	0.12	—
Random effects on melody ID	(Intercept)	22.2	—
	L1onsetIC	2,878.7	—
Residual variance		439.9	—

Note. CI = confidence interval; IC = information content; ID = identity; IOI = interonset interval.

(*Appendices continue*)

Appendix C

Results of Submodel CSTSA Analysis

Perceptual salience is explained in a variety of ways in the current literature, and there is currently no consensus on the correct way to describe or measure it. Dibben (1999) describes salience in relation to pitch register, parallelism, and stability, Collins, Laney, Willis, and Garthwaite (2011) and Huron (2001) in terms of repetition, Prince et al. (2009) in terms of complexity, defined by number of different possible values (i.e., different pitches or different rhythmic durations), and Lerdahl (1989), as a set of conditions combining pitch register, timing, timbre, attack, note density, motivic content, and grouping. Outside of music, salience is defined by predictability and uncertainty (Esber & Haselgrove, 2011); there are two possibilities: predictable content becomes more (i.e., a cue becomes salient if it predicts a reward) or less (i.e., new information is more interesting) salient. Whereas we interpret larger CSTSA model coefficients to reflect more salient predictors, here we test the hypothesis that melodies with high pitch predictability (expectancy) and low-onset predictability (PU) have larger pitch coefficients than onset coefficients, and vice versa. To do so, four submodels of each of the three CSTSA

models optimized in the main experiment were created, one for each category (PP, PU, UP, and UU) to compare coefficients between models. Details of these models can be found in Tables C1–C3. A linear multiple regression model with stimulus predictability, lag type (pitch, onset), and rating type (expectancy, arousal, and valence) predicting the coefficients of these CSTSA models revealed no significant effects, $F(3, 168) = .50$, $p = .67$, $F(1, 170) = 2.23$, $p = .13$, $F(2, 169) = .51$, $p = .59$, respectively. There were also no interactions between category and lag type, $F(7, 164) = .79$, $p = .59$. Although there was no statistically significant effect, we observe that the sum of lags of onset IC were consistently larger than the sum of lags of pitch IC for all categories of stimulus predictability for arousal and valence models, whereas the sum of lags of pitch IC were slightly larger than the sum of lags of onset IC in the expectancy model. In conclusion, our hypothesis was not supported, as salience does not seem to be related to expectancy. However, this study was not designed to investigate this question, which would be interesting to explore in a future study.

Table C1

Coefficients of Submodels for Expectancy Ratings

Model component	Coefficient	PP	PU	UP	UU
Fixed effects	Intercept	0.393	0.399	0.235	0.204
	Time	0.002	0.001	0.003	0.001
	Musicianship	0.035	0.031	0.024	0.031
	Pitch	−0.001	0.002	−0.002	−0.002
	IOI	0.001	0.003	0.004	0.005
	L1ratings	0.777	1.00	1.01	1.03
	L2ratings	0.042	−0.100	−0.141	−0.110
	L4ratings	−0.045	−0.075	−0.075	−0.035
	L5ratings	−0.010	−0.031	0.031	0.021
	L6ratings	0.052	−0.007	0.039	0.008
	L15ratings	0.027	0.013	0.019	0.009
	PitchIC	−0.192	−0.197	−0.142	−0.549
	L1pitchIC	0.197	0.461	0.604	0.842
	L7pitchIC	0.242	0.210	−0.021	0.008
	L8pitchIC	−0.005	−0.169	0.030	−0.087
	OnsetIC	−0.756	−1.26	−0.316	−0.769
	L1onsetIC	1.15	1.50	0.687	0.277
	L2onsetIC	−0.258	−0.469	−0.231	0.078
	L10onsetIC	−0.530	−0.169	−0.038	−0.153
	L12onsetIC	0.188	0.034	0.058	0.235
Random effects	Participant—Intercept	0.002	0.000	0.002	0.000
	Participant—l1pitchIC	0.014	0.003	0.000	0.000
	Melody ID—Intercept	0.105	0.005	0.010	0.006
	Melody ID—l1pitchIC	0.078	0.174	0.337	0.178
Residual variance		0.455	0.397	0.536	0.319

Note. IC = information content; PP = predictable pitch and predictable onset; PU = predictable pitch and unpredictable onset; UP = unpredictable pitch and predictable onset; UU = unpredictable pitch and unpredictable onset; ID = identity; IOI = interonset interval.

(*Appendices continue*)

Table C2

Coefficients of Submodels for Arousal Ratings

Model component	Coefficients	PP	PU	UP	UU
Fixed effects	Intercept	−5.66	0.77	−4.68	−3.45
	Time	0.25	0.04	0.24	0.05
	Pitch	−0.02	−0.02	−0.10	−0.00
	IOI	−0.03	0.00	−0.12	−0.00
	L1ratings	0.87	0.96	0.89	0.95
	L3ratings	0.02	0.01	0.01	−0.00
	L5ratings	−0.05	−0.08	−0.03	−0.02
	L6ratings	0.03	0.04	0.04	0.01
	L15ratings	0.05	0.01	0.02	0.01
	PitchIC	−18.3	−12.7	−11.6	−20.5
	L1pitchIC	19.8	8.85	14.5	24.1
	L6pitchIC	9.35	1.42	−1.65	0.32
	L7pitchIC	−2.64	3.32	8.12	2.05
	L8pitchIC	−0.20	−3.68	−3.71	−4.12
	L10pitchIC	6.03	0.66	2.15	0.08
	L11pitchIC	−10.0	−6.78	2.50	−1.29
	L12pitchIC	6.54	8.64	−2.35	1.75
	L13pitchIC	0.08	−7.88	3.12	−0.45
	L15pitchIC	3.69	0.04	2.14	5.71
	OnsetIC	−24.7	−21.1	3.90	−7.01
	L1onsetIC	78.4	91.4	123.5	17.4
	L3onsetIC	10.2	1.84	10.5	4.73
	L4onsetIC	−15.0	−7.82	−7.92	−2.71
	L7onsetIC	−3.79	7.51	0.01	−1.55
	L10onsetIC	−24.2	−1.00	−6.14	−1.82
	L12onsetIC	13.5	−3.32	5.03	6.53
	L13onsetIC	6.51	11.9	1.76	0.47
	L14onsetIC	−3.12	−6.33	−4.64	−0.81
	L15onsetIC	1.30	1.28	5.12	2.40
Random effects	Participant—Intercept	0.44	1.17	0.12	0.72
	Participant—l1onsetIC	0.51	0.08	0.24	1.10
	Melody ID—Intercept	35.3	8.23	56.4	3.90
	Melody ID—l1onsetIC	2,443.5	3,846.0	11,190.0	447.2
Residual variance		368.1	225.3	323.7	186.2

Note. CI = confidence interval; IC = information content; PP = predictable pitch and predictable onset; PU = predictable pitch and unpredictable onset; UP = unpredictable pitch and predictable onset; UU = unpredictable pitch and unpredictable onset; ID = identity; IOI = interonset interval.

(*Appendices continue*)

Table C3
Coefficients of Submodels for Valence Ratings

Model component	Coefficients	PP	PU	UP	UU
Fixed effects	Intercept	−2.42	9.68	1.35	−1.88
	Time	0.16	0.00	0.17	0.05
	Pitch	−0.06	−0.11	−0.13	−0.00
	IOI	0.17	0.14	0.23	0.14
	L1ratings	0.90	0.94	0.88	0.92
	L3ratings	−0.00	−0.01	−0.05	−0.00
	L4ratings	−0.02	−0.01	−0.05	−0.02
	L5ratings	−0.02	−0.04	0.02	−0.01
	L6ratings	0.02	0.02	0.00	0.01
	L7ratings	0.00	0.00	0.03	0.00
	L9ratings	0.01	0.00	0.02	0.00
	L15ratings	0.01	0.00	0.02	0.01
	PitchIC	−8.42	−6.56	−7.56	−14.7
	L1pitchIC	12.9	8.24	6.30	19.05
	L5pitchIC	2.98	−0.13	3.96	3.18
	L8pitchIC	−1.07	−3.18	−3.32	−8.26
	L9pitchIC	3.10	5.09	1.15	5.82
	L11pitchIC	−0.54	−7.44	2.52	−0.53
	L12pitchIC	4.25	1.02	2.46	2.62
	L15pitchIC	1.71	−0.696	2.30	2.92
	OnsetIC	−22.7	−24.0	−4.81	−21.2
	L1onsetIC	49.1	80.7	81.4	6.13
	L3onsetIC	17.6	3.28	0.22	3.28
	L4onsetIC	−13.1	−0.27	−1.83	−2.87
	L10onsetIC	−14.3	−4.66	−6.26	0.21
	L13onsetIC	7.75	6.75	−0.15	−0.94
Random effects	Participant—Intercept	0.03	0.72	0.10	0.20
	Participant—l1onsetIC	0.14	0.00	1.37	1.06
	Melody ID—Intercept	46.5	9.41	49.8	4.26
	Melody ID—l1onsetIC	2,972.0	4,176.0	5,024.3	36.0
Residual variance		519.6	375.9	712.1	251.5

Note. IC = information content; PP = predictable pitch and predictable onset; PU = predictable pitch and unpredictable onset; UP = unpredictable pitch and predictable onset; UU = unpredictable pitch and unpredictable onset; ID = identity; IOI = interonset interval.