

1 **Pupil responses to pitch deviants reflect predictability of melodic sequences**

2 Bianco Roberta<sup>a</sup> \*, Ptasczynski Lena Esther<sup>b</sup>, Omigie Diana<sup>b</sup>

3 <sup>a</sup> Ear Institute, University College London, 332 Grays Inn Road, WC1X 8EE London, United Kingdom

4 <sup>b</sup> Department of Psychology, Goldsmiths, University of London, 8 Lewisham Way, SE14 6NW London,  
5 United Kingdom

6 **\*Correspondence:**

7 Roberta Bianco

8 [r.bianco@ucl.ac.uk](mailto:r.bianco@ucl.ac.uk)

9

10

11 **Keywords: music, melodies, contextual entropy, deviants, pupillometry**

12 **ABSTRACT**

13 Humans automatically detect events that, in deviating from their expectations, may signal prediction  
14 failure and a need to reorient behaviour. The pupil dilation response (PDR) to violations has been  
15 associated with subcortical signals of arousal and prediction resetting. However, it is unclear how the  
16 context in which a deviant occurs affects the size of the PDR. Using ecological musical stimuli that we  
17 characterised using a computational model, we showed that the PDR to pitch deviants is sensitive to  
18 contextual uncertainty (quantified as entropy), whereby the PDR was greater in low than high entropy  
19 contexts. The PDR was also positively correlated with unexpectedness of notes. No effects of music  
20 expertise were found, suggesting a ceiling effect due to enculturation. These results show that the  
21 same sudden environmental change can lead to differing arousal levels depending on contextual  
22 factors, providing evidence for a sensitivity of the PDR to long-term context.

23

24 **Introduction**

25 The experience of surprise is very common in the sensory realm and often triggers automatic changes  
26 in arousal and attentional states that are fundamental to adaptive behaviours. Music is a ubiquitous  
27 and ecological example of a situation where changes in listeners' arousal and attention are  
28 intentionally manipulated. Composers may, for example, modulate the predictability of musical  
29 passages in order to achieve differing levels of tension in a listener. A great deal of empirical work has  
30 shown that surprising sounds are recognised by listeners in an effortless and automatic fashion  
31 (Pearce, 2018). This process is thought to be supported by a mismatch between the current  
32 unexpected input and the implicit expectations made possible by schematic and dynamic knowledge  
33 of stimulus structure (Huron, 2006; Krumhansl, 2015; Tillmann, Bharucha, & Bigand, 2000; Vuust &  
34 Witek, 2014). However, there is still rather little research examining mismatch responses under  
35 different degrees of uncertainty during passive listening.

36 Evidence of listeners experiencing events in music as unexpected comes from studies investigating  
37 behavioural (Marmel, Tillmann, & Delbé, 2010; Omigie, Pearce, & Stewart, 2012; Tillmann & Lebrun-  
38 Guillaud, 2006) and brain responses to less regular musical events (Bianco, Novembre, Keller, Kim, et

39 al., 2016; Carrus, Pearce, & Bhattacharya, 2013; Koelsch, 2016; Koelsch, Gunter, et al., 2002; Maess,  
40 Koelsch, Gunter, & Friederici, 2001; Miranda & Ullman, 2007; Omigie, Pearce, Williamson, & Stewart,  
41 2013; Pearce, Ruiz, Kapasi, Wiggins, & Bhattacharya, 2010). With regard to the former, priming  
42 paradigms have shown that a context allows perceivers to generate implicit expectations for future  
43 events, leading to facilitated processing (i.e., priming) of expected events. With regard to the latter,  
44 violation paradigms have shown increased brain responses to deviant events (out of key notes, or  
45 harmonically incongruent chords) within structured contexts as well as events which are musically  
46 plausible but more improbable in the given context. For example, Omigie et al. (2013) tested brain  
47 responses to melodies whose notes were characterised in terms of their predictability by a model of  
48 auditory expectations (Pearce, 2005). They showed that surprising events (more improbable notes)  
49 within melodies elicited a mismatch response – often termed the mismatch negativity, MMN (Garrido,  
50 Kilner, Stephan, & Friston, 2009; Näätänen, Paavilainen, Rinne, & Alho, 2007). This response decreased  
51 in amplitude for progressively more predictable events, as estimated by a computational model of  
52 melodic expectation. A similar parametric sensitivity to note unexpectedness has since also been  
53 observed in subcortical regions like the anterior cingulate and insula (Omigie et al., 2019). Moreover,  
54 sensitivity to music structure violation seems to emerge in all members of the general population that  
55 have had sufficient exposure to a given musical system (Bigand & Poulin-Charronnat, 2006; Pearce,  
56 2018; Rohrmeier, Rebuschat, & Cross, 2011), and this sensitivity is modulated by pre-existing  
57 schematic knowledge of music, as reflected in acquired levels of musical expertise (Fujioka, Trainor,  
58 Ross, Kakigi, & Pantev, 2004; Koelsch, Schmidt, & Kansok, 2002a; Tervaniemi, 2009; Vuust, Brattico,  
59 Seppänen, Näätänen, & Tervaniemi, 2012).

60 According to theoretical and empirical work framing perception in the context of predictive processing,  
61 the experience of surprise may be modulated by the predictability of a stimulus structure as it unfolds  
62 (Clark, 2013; Dean & Pearce, 2016; Friston, 2005; Ross & Hansen, 2016). Random or high entropic  
63 stimuli hinder the possibility of making accurate predictions about possible upcoming events, whilst  
64 stimuli characterized by familiarity or regular statistics will enable relatively precise predictions by  
65 permitting the assignment of high probability to a few possible continuations. Perceptually, it has been  
66 shown that listeners indeed experience high-entropic musical contexts with greater uncertainty  
67 compared to low entropy ones (Hansen & Pearce, 2014). Moreover, previous work has shown that  
68 neurophysiological signatures associated with auditory surprise display larger responses to a given  
69 deviant event when it is embedded in a low rather than high entropic context (Dean & Pearce, 2016;  
70 Garrido, Sahani, & Dolan, 2013; Hsu, Le Bars, Hamalainen, & Waszak, 2015; Ricardo Quiroga-Martinez,  
71 2018; Rubin, Ulanovsky, Nelken, & Tishby, 2016; Southwell & Chait, 2018). Therefore, research  
72 suggests that to understand whether and how surprising events modulate arousal and re-orient  
73 behaviours, the statistics of the proximal context must be considered.

74 A vast literature has used pupil dilation response (PDR) as a general marker of arousal, selective  
75 attention, and surprise (Aston-Jones & Cohen, 2005; Sara, 2009). Pupil dilation is associated with the  
76 locus coeruleus-norepinephrine (LCN) system (Laeng, Sirois, & Gredeback, 2012; Widmann, Schröger,  
77 & Wetzel, 2018), the activation of which results in wide spread norepinephrine release in the brain.  
78 Increase of the PDR has been extensively reported in response to violation of expectations or  
79 surprising/salient events in distracted listeners (Damsma & van Rijn, 2017; Liao, Yoneya, Kidani,  
80 Kashino, & Furukawa, 2016; Wetzel, Buttellmann, Schieler, & Widmann, 2016; Zhao et al., 2019), and  
81 when deviants are presented below participants' perceptual threshold (Fink, Hurley, Geng, & Janata,  
82 2018). Furthermore, a relationship between PDR and continuous ratings of surprisal in music  
83 (measured via a continuous slider) has been reported during passive listening (Liao, Yoneya, Kashino,

84 & Furukawa, 2018). Other work has specifically associated the PDR to violations of statistical  
85 regularities in the sensory input even when these violations are behaviourally irrelevant (Alamia,  
86 VanRullen, Pasqualotto, Mouraux, & Zenon, 2019; Zhao et al., 2019). These results provide supporting  
87 evidence for a role of norepinephrine in the tracking of abrupt deviations from the current predictive  
88 model of the world, and as a signal of prediction resetting that enables the discovery of new  
89 information (Dayan & Yu, 2006).

90 Music has the ability to play with our expectations, hence manipulating our arousal and emotions in  
91 an automatic fashion (Laeng, Eidet, Sulutvedt, & Panksepp, 2016; Meyer, 2001; Zatorre & Salimpoor,  
92 2013). Abrupt changes in register, texture and tonality are all examples of instances where the listener  
93 may have to reset or potentially abandon current models about the unfolding music. Such changes  
94 may however appear less surprising if embedded in high entropy contexts. Here, we use music as an  
95 ecological setting with which to study how the PDR to deviant musical events is modulated by structure  
96 of the stimulus context, specifically by its entropy. The growing field of computational musicology  
97 means that information in melodies can be statistically estimated. One particular model of auditory  
98 expectations – the Information Dynamics of Music, IDyOM (Pearce, 2005) – has been shown to model  
99 listeners’ experience of surprise and uncertainty. This unsupervised Markov model learns and  
100 estimates the conditional probability of each subsequent note in a melody based on a corpus on which  
101 it has been trained (long-term sub-model; extra-opus-learning) and the given melody as it unfolds  
102 (short-term sub-model; intra-opus learning). The model outputs information content and entropy  
103 values, which, respectively, reflect the experienced unexpectedness of a certain note after its onset  
104 and the experienced uncertainty in precisely predicting a subsequent note based on the preceding  
105 pitch probabilities.

106 We created novel melodies (Figure 1) that adhered to the principles guiding Western tonal melodic  
107 structure. We then created shuffled versions of these melodies to create stimuli that were higher in  
108 entropy albeit matched for pitch range, content and tonal space. The information theoretic properties  
109 of all stimuli were estimated using IDyOM (Pearce, 2005). Listeners were presented with these  
110 melodies either in their standard form or with a pitch deviant whilst PDR was measured. Participants  
111 were not informed about the presence of the pitch deviants, but asked to rate the unexpectedness of  
112 last note of each melody. We expected a larger PDR to deviant notes that are embedded in low rather  
113 than high entropy contexts, and that are higher in their unexpectedness – information content value  
114 – as estimated by the computational model. Also, we expected entropy of the melodies to predict  
115 subjective ratings of stimulus unexpectedness (Hansen & Pearce, 2014). Finally, based on evidence of  
116 greater brain response to musical violations in musicians than non-musicians (Vuust, Brattico,  
117 Seppänen, Näätänen, & Tervaniemi, 2012b), presumably reflecting expertise-related enhancement in  
118 the accuracy of predictive models, differences due to musical expertise (Müllensiefen et al., 2014) were  
119 also investigated.

## 120 **Methods**

### 121 *Participants*

122 Forty-seven participants (age:  $M=26.19$ ,  $SD=6.24$ , min.=20, max.=52, 68% female, representing 15  
123 nationalities) took part in the study. The sample scored relatively high on the general musical  
124 sophistication index, GoldMSI (Müllensiefen et al., 2014), with  $M=87.40$ ,  $SD=25.24$ , min.=32,  
125 max.=120. A big sample size was chosen based on a previous experiment using musical stimuli (Laeng  
126 et al., 2016) and to ensure statistical power. As a post-hoc confirmation of the adequacy of our sample-

127 size, we quantified effect sizes of the difference between PDR to deviants in high vs. low entropy  
128 contexts. The power analysis was conducted in the G\*Power software package (version 3.1.9.2) with  
129 the setting  $p = .05$  and  $N = 42$  and confirmed that our sample size was adequate ( $1 - \beta > .9$ ). Two  
130 participants reported ophthalmologic concerns or surgery prior to the experiment but were not  
131 excluded from participation as the pupil dilates even in blindsight participants (Weiskrantz, Cowey, &  
132 Barbur, 1999). Technical problems occurred during the recording of four participants, who were  
133 therefore excluded from the analysis. One subject was further excluded as blink gaps were too large  
134 to be interpolated. In sum, forty-two participants' data were analysed.

135 Ethical approval for this study was granted by the Research Ethics Committee of Goldsmiths, University  
136 of London. Participants were instructed as to the purpose of the study, and consented to participate  
137 (written informed consent). Participation was remunerated with 5 pounds.

### 138 *Stimuli*

139 One-hundred-twenty melodies were used in this study (thirty originally composed, thirty matched  
140 'shuffled' versions, and sixty corresponding versions with a deviant tone always as the 13<sup>th</sup> note). The  
141 melodic sequences comprised 20 tones, were 5 seconds long, isochronous (with an inter-onset-interval  
142 of 250 ms, 20/4 bar with 240 bpm), and had constant intensity and timbre (MIDI generated piano  
143 timbre).

144 The corpus of melodies was composed according to the principles of Western-tonal music and using  
145 all tones of the chromatic scale. Ambitus and tonal space varied across melodies. Interval size did not  
146 exceed a perfect fifth (Narmour, 2015) between adjacent notes. Thus, the original melodies were  
147 characterized by a smooth contour. To generate matched melodies that controlled for potential biases  
148 such as tonal space, pitch class, frequency range or ambitus, high entropy melodies were created from  
149 the original melodies by randomizing the order of constituent notes using MIDI processing tools (Eerola  
150 & Toiviainen, 2004). Our manipulation of entropy, whereby notes in original melodies were randomly  
151 shuffled without constraint, necessarily resulted in the mean absolute interval size being greater in  
152 high entropy than in low entropy melodies [L vs. H:  $t(58) = -2.644$ ,  $p\text{-value} = .01$ ]. Corpus studies of  
153 western tonal music show that large interval sizes are much less common than small ones (Huron,  
154 2001). Thus, we anticipated that the presence of such large intervals would lead to the higher entropy  
155 levels desired for the high entropy condition.

156 Deviant notes were inserted in all 60 melodies (original and shuffled version) at the onset of the 13<sup>th</sup>  
157 note (3000ms on the salient onbeat) in order to create the corresponding set of *deviant melodies* (Fig.  
158 1). The deviant note was integrated into the second half of the melody in order to allow the  
159 establishment of an expectation-forming context before its occurrence. To ensure that differences in  
160 the PDR to the deviant notes between low and high entropy condition were not attributable to  
161 difference in the just preceding event, but to the context, the interval size between the deviant and  
162 the preceding 12th note was the same in matched pairs of high and low entropy melodies (see Fig. 1).  
163 However, interval size of the deviant varied between maj7 up/down, min9 up/down and aug11  
164 up/down as those intervals sound particularly unusual within a melodic progression. Larger interval  
165 size of the deviant was assigned to matched pairs with lower entropy differences. This allowed for a  
166 variety of deviant interval sizes (a range of unexpectedness) while ensuring salience of deviants even  
167 in melodies showing relatively smaller entropy differences.

168 Entropy values were assigned by the IDyOM model to each note of the melodies. The IDyOM model  
169 considered one pitch viewpoint, namely 'cpitch', whereby chromatic notes count up and down from

170 the middle pitch number (C=60) (Pearce, 2005). Through a process of unsupervised learning, the model  
171 was trained on a corpus (903 Western tonal melodies) comprising songs and ballads from Nova Scotia  
172 in Canada, German folk songs from the Essen Folk Song Collection and chorale melodies harmonised  
173 by J.S. Bach. The probability of each note of the stimulus used here were then estimated based on a  
174 combination of the training set's statistics and those of the given melody at hand. The model outputs  
175 information content and entropy values. In mathematical terms, information content is inversely  
176 proportional to the probability of an event  $x_i$ , with  $IC(x_i) = -\log_2 p(x_i)$  (MacKay, 2003), while the  
177 maximum entropy is reached when all potential events  $x_i$  are equally probable, with  $p(x_i) = 1/n$ , where  
178  $n$  equals the number of stimuli. In psychological terms, information content represents how surprising  
179 each subsequent note is based on its fit to the prior context (Pearce & Wiggins, 2006). In contrast,  
180 entropy refers to the anticipatory difficulty in precisely predicting a subsequent note. Mean entropy  
181 values, obtained by taking the mean entropy of all notes, were used to characterise the predictability  
182 of each melody. Given our manipulation, mean entropy was largely, but not entirely, explained by the  
183 mean interval size of melodies ( $\rho = .474$ ,  $p < .001$ ). These values were then used to predict subjective  
184 inferred uncertainty (measured as unexpectedness of last note) about each melody, and to categorise  
185 melodies into low and high entropy groups. By controlling that the interval preceding the deviant note  
186 was identical across high and low entropy condition, we predicted that a greater PDR to deviant notes  
187 in low than high entropy melodies should be attributable to the difference in the context.

#### 188 *Procedure*

189 The experiment was presented using the Experiment Builder Software and pupil diameter was  
190 recorded using EyeLink 1000 eye-tracker at a 250Hz sampling rate (SR Research, [www.sr-research.com/experiment-builder](http://www.sr-research.com/experiment-builder)). Prior to the data acquisition, a three-dot calibration was  
191 conducted to ensure adequate gaze measurements. Participants were further allowed to adjust the  
192 sound volume to a comfortable level and were asked to reduce head movements to a minimum  
193 throughout the recording session. As no differences were anticipated between the left and right pupil,  
194 the left pupil was recorded in ten and the right pupil in 32 participants depending on the participants'  
195 dominant eye. To reduce motion artefacts, the head was stabilized using the SR Research Head Support  
196 chinrest placed 50 cm from the presentation screen.  
197

198 During the experiment, the 120 melodies were presented binaurally through headphones in a  
199 randomized order. Each trial was triggered by the experimenter on the control computer when the  
200 fixation was stable at less than two arbitrary gaze units away from the fixation point. When recording  
201 was enabled, a white fixation dot on the grey screen turned black to prepare participants for the onset  
202 of the melody. The fixation cross was displayed for 7 seconds after stimulus onset. Each trial was  
203 preceded by a 400 ms baseline period and followed by a 2000 ms post stimulus offset. The melody was  
204 5000 ms long and participants were instructed not to blink or move but to fixate during that whole  
205 period. Participants were allowed to take breaks to avoid fatigue effects at their convenience. A re-  
206 calibration procedure using the 3-dot-calibration was applied after each break.

207 Participants were instructed to carefully listen to the melody while fixating on the fixation point in the  
208 centre of the screen throughout the recording period. They were not informed about the deviant  
209 manipulation. After each trial, participants rated the final note on a Likert-scale ranging from 1 (not at  
210 all unpredictable) to 7 (extremely unpredictable) in a forced-choice task on the presentation screen.  
211 Data on the subjects' musical expertise and sociodemographic background was collected at the end of  
212 each experiment using the GoldMSI (Müllensiefen et al., 2014). The whole study lasted approximately  
213 one hour.

## 214 *Data pre-processing*

215 Blinks were identified and removed from the signal using MATLAB R2017b. These were characterized  
216 by a rapid decline towards zero from blink onset, and a rapid rise back to the regular value at blink  
217 offset. 100ms of the signal was removed before and after the missing data points (Troncoso, Macknik,  
218 & Martinez-Conde, 2008) and missing data were interpolated: four equally spaced time points were  
219 used to generate a cubic spline fit to the missing time points between blink onset ( $t_2$ ) and blink offset  
220 ( $t_3$ ) of the unsmoothed signal, with  $t_1 = t_2 - t_3 + t_2$  and  $t_4 = t_3 - t_2 + t_3$  (Mathôt, Fabius, Van Heusden, & Van  
221 der Stigchel, 2018). Trials containing more than 15% missing data were excluded from the analysis ( $M =$   
222  $.3$ ,  $SD = .9$  trials across subjects). Data were cleaned of artefacts using Hampel filtering (median filtered  
223 data; Pearson, Neuvo, Astola, & Gabbouj, 2016), and smoothed using a Savitzky-Golay filter of  
224 polynomial order 1 over the entire trial epoch. Finally, data were z-scored, and then baseline-corrected  
225 by subtracting the median pupil size of 400 ms baseline before melody onset. To analyse the time-  
226 window after deviant onset, data were baselined from 400 ms before deviant onset. The normalized  
227 pupil diameter was time-domain-averaged across trials of each condition.

## 228 *Experimental design and statistical analysis*

229 We estimated a single time series for each sequence type: low or high entropy with either standard or  
230 deviant note type (*S-Low/D-Low/S-High/D-high*) by averaging across trials and participants. Statistical  
231 analysis was performed using Fieldtrip's cluster-based permutation test (Maris & Oostenveld, 2007),  
232 with a significance threshold at 5% to control family-wise error-rate (FWER). This analysis revealed the  
233 time windows showing significant difference between the compared time series.

234 We first compared responses to deviant notes in the low and high entropy contexts with the  
235 corresponding standard notes using signals recorded across the entire melody duration (baselined 400  
236 ms before melody onset). This ensured that differences between deviant and corresponding standards  
237 were not due to differences in the immediately preceding note. Then, to determine how the deviant  
238 PDR is affected by the entropy of the melodic context, we focussed on the time window from deviant  
239 onset to the end of the melody (3000 to 5000 ms epochs). We thus baselined to 400 ms before deviant  
240 onset, and we tested for an interaction of the deviant and context manipulation: (*D-Low – S-low*) vs.  
241 (*D-High – S-High*). Further, we compared the responses to standard tones in the two contexts (*S-Low*  
242 *– S-High*) to ensure that any differences were not driven simply by the standard notes (the control  
243 conditions).

244 To assess a potential influence of expertise on the PDR to deviants (data between 3000 and 5000 ms  
245 baselined to 400 ms before deviant onset), we computed the mean PDR to deviant trials as *D-Low – S-*  
246 *Low* for deviants in low entropy context, and *D-High – S-High* for deviants in high entropy context.  
247 Participants were split into two groups of musical experts and non-experts based on GoldMSI scores  
248 ( $Mdn=96$ ). An ANOVA with within-subject factor context (low/ high entropy) and expertise as between-  
249 subject factor (expert/non-expert) was computed.

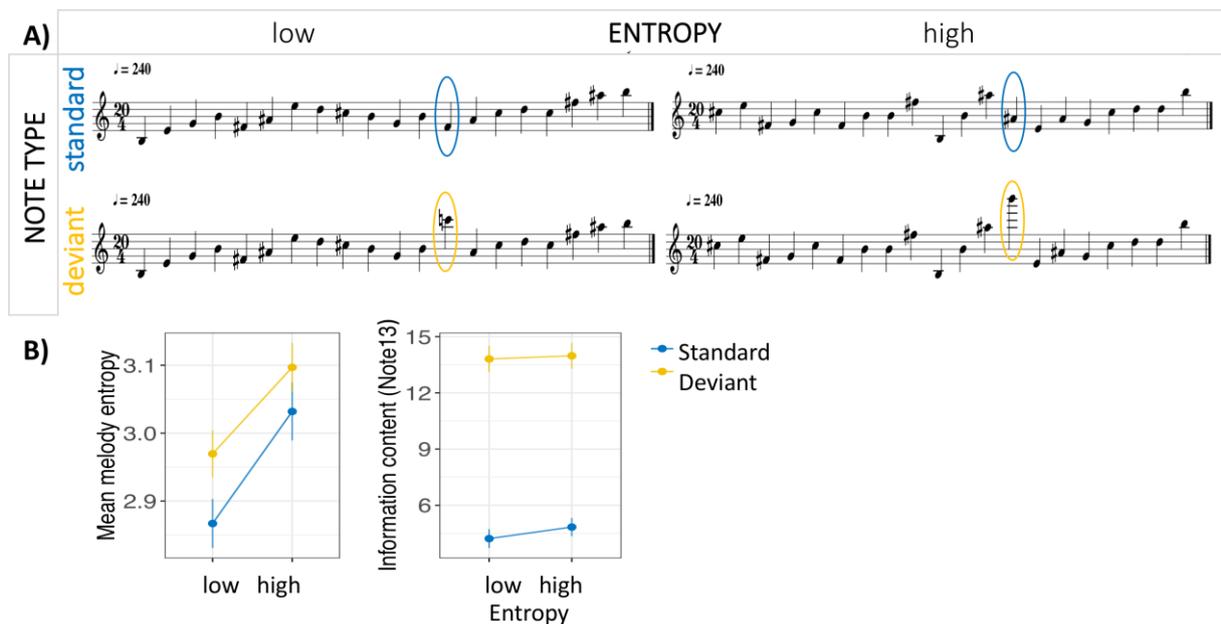
## 250 **Results**

251 *Stimuli characterization.* Analyses were carried out to clarify the nature of all differences in information  
252 theoretic properties of the different stimuli. Figure 1B (left panel) shows the mean entropy values for  
253 all conditions (*D-Low*:  $M = 2.96$ ,  $SD = .18$ ; *D-High*:  $M = 3.09$ ,  $SD = .19$ ; *S-Low*:  $M = 2.86$ ,  $SD = .19$ ; *S-High*:  
254  $M = 3.03$ ,  $SD = .23$ ). An ANOVA with context (low/ high entropy) and note type (deviant / standard) as  
255 between-group factors yielded a main effect of context [ $F(1,116) = 15.20$ ,  $p < .001$ ,  $np2 = .12$ ],

256 indicating higher entropy in *High* than *Low entropy* melodies [ $t(58) = 2.85, p = .006$ ]. A significant main  
 257 effect of note type [ $F(1,116) = 5, p = .027, np2 = .04$ ], was not supported by a further post hoc  
 258 comparison [ $t(58) = -1.641, p = .11$ ]. No interaction was found between the two factors [ $F(1,116) = .25,$   
 259  $p = .617, np2 < .01$ ].

260 Figure 1B (right panel) shows that the unexpectedness of deviant notes, as reflected by information  
 261 content values, was comparable between low and high entropy melodies. An ANOVA with between  
 262 group factors context (low / high entropy) and note type (deviant / standard) yielded a main effect of  
 263 note type [ $F(1,116) = 239.02, p < .001, np2 = .67$ ], a non-significant main effect of context [ $F(1,116) =$   
 264  $.42, p = .517, np2 = .01$ ], and no interaction between the two factors [ $F(1,116) = .00, p = .714, np2 <$   
 265  $.01$ ] – thus indicating higher information content in deviant than standard notes regardless of the  
 266 context in which they were embedded [ $t(58) = -12.321, p < .001$ ]. The similar IC levels of deviant for  
 267 the two entropy conditions may be due to fact that in both low and high entropy contexts, deviants  
 268 were similar in being characterised by very large interval departures away from the melodic contour  
 269 (as opposed to the relatively naturalistic events that occurred in real melodies (Dean & Pearce, 2016).  
 270 Critically, that deviant IC levels are similar for the two entropy conditions, supports our suggestion that  
 271 stimulus context (and not just the IC level an incoming event) has the ability to modulate the PDR.

272 IC was positively predicted by interval size between the 12th and 13th note ( $\rho = .862, p < .001$ ), in  
 273 line with research showing that amongst multiple psychological representations of pitch (e.g., height,  
 274 contour, etc.), interval exerts a major contribution to perception of surprise (Levitin & Tirovolas, 2009;  
 275 Pearce, 2018; Quiroga-martinez et al., 2019).



276

277

278 **Fig. 1. Experimental design.** Participants listened to the melodies of 20 notes and had to rate how  
 279 unexpected they found the last note on a seven-step Likert scale (– 1 equal ‘not at all unpredictable’  
 280 and 7 ‘extremely unpredictable’). The design included two factors: Contextual entropy (low/high) and  
 281 Note type (standard/deviant). Melodies containing deviant notes where 50% of the trials and the  
 282 deviant notes occurred always at the 13<sup>th</sup> position. **B) Characterization of the stimuli.** A model of  
 283 musical expectations was used to characterise the stimuli: Mean entropy was larger for high than low  
 284 entropy melodies regardless of the presence of deviant (yellow) or standard (blue) tones. Mean

285 information content of deviants (yellow) was larger than standard (blue) tones regardless of the  
286 entropy of the context.

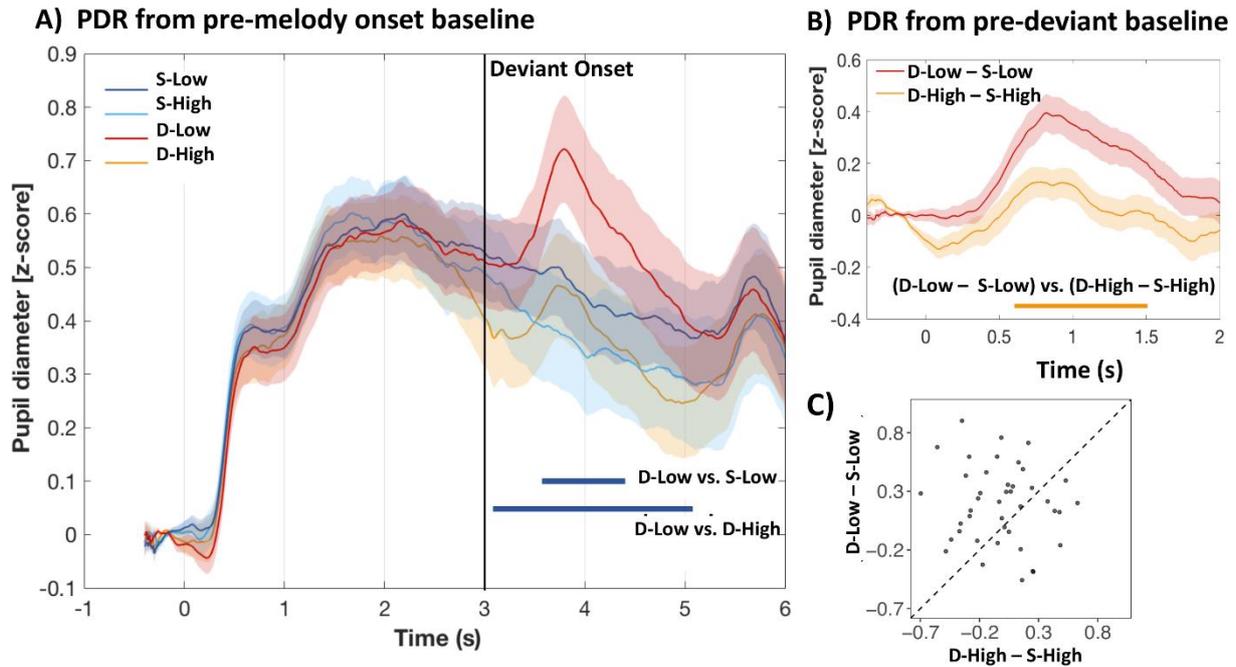
287 *PDR characterization.* Figure 2A shows the time course of the PDR across conditions (*S-Low* = standard  
288 low entropy, *S-High* = standard high entropy, *D-Low* = deviant low entropy, *D-High* = deviant high  
289 entropy) baselined 400 ms before melody onset. A comparison between *S-Low* and *S-High* showed no  
290 difference in the PDR as a function of entropy levels of the melody, whilst the PDR to deviants (*D-Low*  
291 vs. *D-High*) was greater in predictable than unpredictable contexts (diverging at .56 s from deviant  
292 onset). The response to deviants in the predictable contexts was greater than the relative standard  
293 condition (*D-Low* vs. *S-Low*:  $p = .029$ ), significantly diverging from *S-Low* between 3.57 and 5.64 s after  
294 melody onset. Conversely, the response to deviants in unpredictable contexts did not differ from the  
295 relative standard condition (*D-High* vs. *S-High*), despite their high information content (see Figure 1B,  
296 middle panel).

297 Figure 2B shows the PDR to deviants baselined 400 ms before deviant onset. We show the conditions  
298 *D-Low* and *D-High* following subtraction of the relative standard conditions. The comparison between  
299 the two time-courses confirmed that *D-Low* evoked a larger response than *D-High* starting at .59 s from  
300 deviant onset ( $p = .007$ ) and ending at 1.5 s. Intrinsic noise in the baseline may explain the very early  
301 divergence between the curves, which however was not significant. Importantly, this pattern was  
302 observed in 67% of the participants (Figure 2C).

303 We run a post hoc analysis (Figure 3B) to investigate the relationship between the PDR related to the  
304 13<sup>th</sup> note (including both standard and deviant notes) and associated information content for low and  
305 high entropy melodies. For each subject and for each melody, the average PDR to the 13<sup>th</sup> note was  
306 computed from the pre-deviant baseline (as in Figure 2B). A linear model with the factor context (Low  
307 /High) and information content as continuous variable was run to predict the PDR. This analysis yielded  
308 no main effect of context [ $F(1,116) = .005$ ,  $p = .941$ ,  $np2 < .001$ ] and a main effect of information  
309 content [ $F(1,116) = 4.911$ ,  $p = .029$ ,  $np2 = .041$ ] and a no significant interaction [ $F(1,116) = 3.274$ ,  $p =$   
310  $.073$ ,  $np2 = .027$ ]. This suggests that the PDR is sensitive to a large range of unexpectedness levels.

311 We further investigated potential influences of expertise on the mean PDR to deviants (computed as  
312 *D-Low* – *S-Low* for deviants in low entropy contexts, and *D-High* – *S-High* for deviants in high entropy  
313 contexts on the data between 3000 and 5000 ms baselined to 400 ms before deviant onset as in Figure  
314 2B). The ANOVA examining the effect of musical expertise on the mean PDR to deviants yielded a main  
315 effect of context [ $F(1,40) = 5.09$ ,  $p = .03$ ,  $np2 = .11$ ]. The main effect of expertise was not significant  
316 [ $F(1,40) = 3.72$ ,  $p = .061$ ,  $np2 = .09$ ], and no interaction between the two factors was seen [ $F(1,40) =$   
317  $2.18$ ,  $p = .147$ ,  $np2 = .05$ ]. This confirmed that the mean PDR to deviants was greater in low than high  
318 entropy contexts although no considerable difference between experts and non-experts was observed.

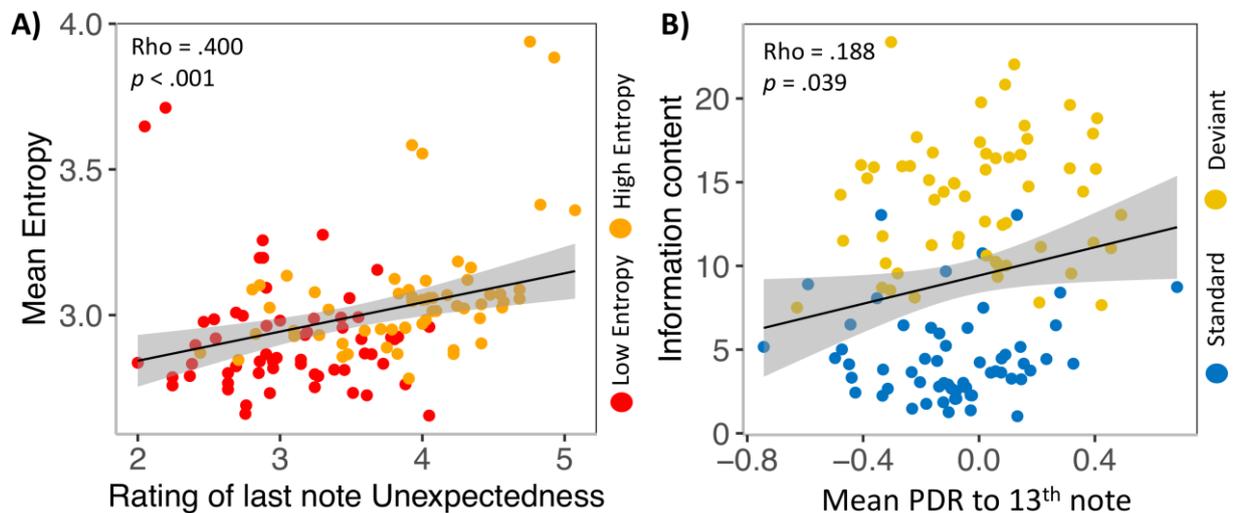
319 Finally, we showed that the model reliably predicted subjective uncertainty levels (inferred entropy)  
320 of the melody progressions (Hansen & Pearce, 2014) (Figure 3A). The measure of the unexpectedness  
321 of the last note was collected after participants listened to each melody. We found that the mean  
322 ratings for each melody strongly correlated with the information content of the last note of the melody  
323 ( $\rho = .355$ ,  $p < .001$ ), and with the mean IDyOM entropy values for that melody ( $\rho = .400$ ,  $p < .001$ ).  
324 This analysis validated our categorization of melodies based on the IDyOM output.



325

326 **Fig. 2. A)** The PDR for all conditions from melody onset (*S-Low* = standard low entropy, *S-High* =  
 327 standard high entropy, *D-Low* = deviant low entropy, *D-High* = deviant high entropy). The PDR to  
 328 deviants compared to standard tones (*D-Low* vs. *S-Low*) in low entropy contexts diverged between  
 329 3.57 and 5.64 s from melody onset (.57 s from deviant onset), but did not differ in high entropy  
 330 contexts. Also, the PDR to deviants was greater in low than in high entropy contexts (*D-Low* vs. *D-High*)  
 331 (diverging at 3.056 s from melody onset), but there was no significant context-dependent difference  
 332 between the standard tones (*S-Low* and *S-High*). Shaded regions around the curves represent standard  
 333 error in the mean estimated with bootstrap resampling (1000 iterations; with replacement). **B)**  
 334 Interaction effect of deviant and context entropy on the PDR after deviant onset. The difference  
 335 between the PDR to deviant and standard tones was greater in low than in high entropy contexts. This  
 336 effect emerged .59 after deviant onset and ended at 1.5 s. **C)** The relationship between the PDR for *D-*  
 337 *High* – *S-High* and *D-Low* – *S-Low* conditions. Each data point represents an individual participant. Dots  
 338 above the diagonal reveal that 67% of participants showed a greater PDR to deviants in low compared  
 339 with high entropy contexts.

340



341

342

343 **Fig. 3. A)** The stimulus entropy was computed by the IDyOM model, and validated by participants' self-  
344 reports about the overall unexpectedness of each melody. Each dot represents one of the 120  
345 melodies. Red and orange dots represent each of the low and high entropy melodies, respectively ( $N$   
346 = 120). Shading represents s.e.m. **B)** The unexpectedness of the 13<sup>th</sup> notes (deviants and standards)  
347 was estimated by the IDyOM model and showed a positive correlation (Spearman) with the mean  
348 evoked PDR. Blue and yellow dots represent each of the standard and deviant notes. Shading  
349 represents s.e.m.

## 350 **Discussion**

351 We report that pupil dilation response (PDR) to behaviourally irrelevant deviants occurs when deviants  
352 are embedded in predictable rather than unpredictable melodies. We showed that the amplitude of  
353 the response is predicted by the information content (or unexpectedness) of the musical deviants. We  
354 also replicate the previous finding that listeners' experience of uncertainty is predicted by the entropy  
355 of the music (Hansen & Pearce, 2014). These results show that the same sudden environmental change  
356 leads to differing levels of arousal depending on whether it occurs in low or high states of uncertainty.  
357 Our results suggest that the more stable predictions formed in predictable rather than unpredictable  
358 contexts may be more abruptly violated by surprising events, possibly leading to greater changes in  
359 the listeners' arousal state.

360 The observed modulatory effect of context predictability on the PDR to deviants is consistent with a  
361 body of electrophysiological work showing context effects on mismatch like responses at the cortical  
362 level (Garrido et al., 2013; Quiroga-Martinez et al., 2019; Southwell & Chait, 2018). Here we show a  
363 similar pattern but in autonomic markers of arousal, as reflected by the PDR. Pupil response is thought  
364 to be driven by norepinephrine activity in the locus coeruleus (LC) (Joshi, Li, Kalwani, & Gold, 2016).  
365 This is a key subcortical nucleus which widely connects to the brain (Sara, 2009) to signal unexpected  
366 and abrupt contextual changes (Alamia et al., 2019; Damsma & van Rijn, 2017; Zhao et al., 2019). It has  
367 been hypothesised (Zhao et al., 2019) that MMN-related brain systems (Garrido et al., 2009; Hsu et al.,  
368 2015; Southwell & Chait, 2018) may trigger norepinephrine-mediated updating or interruption of  
369 ongoing top-down expectations. In line with this hypothesis, Alamia et al. (2019) have shown a  
370 correlation between pupillary response and MMN-like response evoked by violations of expectations,  
371 providing evidence of a link between the sources of these two responses. Top-down expectations  
372 about unfolding sensory signals have been associated with the temporo-frontal network in music-  
373 violation paradigms, which is thought to link present and past information to generate predictions  
374 about forthcoming events (Bianco, Novembre, Keller, Seung-Goo, et al., 2016; Koelsch, Gunter, et al.,  
375 2002; Tillmann, Janata, & Bharucha, 2003). Based on this existing evidence and in line with a model  
376 resetting hypothesis, our results suggest that listeners generate stronger predictive models (in the  
377 predictable melodies). These may be supported by temporo-frontal cortical regions, and require a  
378 greater signal (greater PDR) to be interrupted.

379 Albeit indirectly, our results also establish a link between subcortical and cortical activity in response  
380 to unexpected events under different states of uncertainty. Increased MMN response under low states  
381 of uncertainty (Quiroga-Martinez et al., 2019; Southwell & Chait, 2018) are replicated in the pupil  
382 response, reflecting a general increase in arousal. A possible interpretation, in line with the predictive  
383 coding theory (Friston, 2005), is that more stable expectations, representative of a strong predictive  
384 model, are reflected in precision-weighting of the prediction error, and hence a stronger response when

385 the input mismatches the current predictions. Conversely, in high entropic contexts predictive models  
386 are weak, and the prediction error attenuated. This mirroring pattern between cortical (MMN) and  
387 subcortical (as reflected by the PDR) responses may have important behavioural advantages.  
388 Specifically, strong predictive models can suddenly be abandoned when they are revealed to be  
389 erroneous, thus allowing speedy reorienting behaviours and quicker engagement with new potentially  
390 relevant stimuli. One limitation with regard to this possible interpretation is the nature of the deviant  
391 events used here, whereby the deviating event was a single note that did not lead to any long-lasting  
392 changes in the statistics of the unfolding sequence. Further studies combining cortical and pupil  
393 response measurements in continuously changing stimuli are necessary to corroborate our working  
394 hypothesis.

395 The absence of difference between high and low entropic contexts in the sustained pupil response (in  
396 the first half of the melodies) suggests that the pupil is relatively unresponsive to slowly unfolding  
397 stimulus structures, at least when listeners are not required to actively track them (Alamia et al., 2019;  
398 Zhao et al., 2019). Whilst cortical responses have been shown to be sensitive to the statistics of the  
399 unfolding stimulus structure (Barascud, Pearce, Griffiths, Friston, & Chait, 2016; Sohoglu & Chait, 2016;  
400 Southwell et al., 2017), subcortical responses may be less sensitive. This suggests they may be more  
401 vulnerable to stimulus properties and tasks demands (Zhao et al., 2019).

402 We also found that the PDR mismatch response is positively predicted by unexpectedness of incoming  
403 notes (in line with electrophysiological studies (Omigie et al., 2013; Quiroga-martinez et al., 2019), but  
404 it seems not to be modulated by degree of musical expertise. Larger MMN responses have been shown  
405 for musicians in a range of studies examining electrophysiological correlates of expectancy violation in  
406 music (Koelsch, Schmidt, & Kansok, 2002b; Oechslin, Van De Ville, Lazeyras, Hauert, & James, 2013;  
407 Tervaniemi, Tupala, & Brattico, 2012). One possibility is that this reflects a ceiling effect whereby the  
408 rather salient deviants used here were relatively easy to detect (given their large interval departures  
409 away from the melodic contour). Future studies examining the PDR to more subtle differences in  
410 musical structure may be expected to show similar expertise effects to those reported in previous  
411 studies.

412 Finally, our results provide evidence of music's usefulness in investigating the neural mechanisms  
413 underlying processing of stimuli statistical properties in a common, highly structured, and ecologically  
414 valid type of auditory stimulus, as music. While our work here focused on pitch expectations, previous  
415 studies have shown that music-induced temporal expectations are also tracked by the PDR (Damsma  
416 & van Rijn, 2017). Future experiments could address, for example, how introducing rhythm to  
417 concurrent melodic lines may affect the PDR to unexpected events. Similarly, that music induced chills  
418 – high arousal physiological responses associated with subjective pleasure – are associated with  
419 increased PDR (Laeng et al., 2016) suggest a usefulness of music in examining the relationship between  
420 stimulus information theoretic properties and reward processing. Whilst the IDyOM model used here  
421 is only able to deal with monophonic MIDI music, the development of the model for polyphonic music  
422 is underway (<https://psyarxiv.com/wgjyv/>) and it is expected that the approach we take here will be  
423 beneficial in a wider range of contexts in future years. By showing that predictive uncertainty can be  
424 used to modulate prediction-error related arousal, our findings have implications for understanding  
425 the variety of forms listeners' aesthetic appreciation of music may take. However, considering, more  
426 generally, the tight coupling between the error-related norepinephrine system and the reward-seeking  
427 dopaminergic pathway (Laeng et al., 2016; Xing, Li, & Gao, 2016; Zatorre & Salimpoor, 2013), our  
428 results emphasize that measuring the PDR may be useful for investigating the reward value of  
429 information across a range of modalities and domains.

430 In sum, we show that pupillometry in the auditory domain can reliably track the effect of context  
431 uncertainty on responses to sudden environmental change and independently from overt behavioural  
432 responses. Given the tight interplay between cortical and subcortical mechanisms involved in precision  
433 weighted anticipatory processing, a first milestone is set towards the non-invasive quantification of  
434 related arousal responses.

435

#### 436 **Competing Interests Statement**

437 The author(s) declares no competing interests.

#### 438 **Author Contributions**

439 **RB** conceived the experiments; analysed the bulk of the data; wrote the manuscript.

440 **EP** conceived and performed the experiments; contributed to analysis; wrote the manuscript; secured  
441 the funding.

442 **DO** conceived the experiments; contributed to analysis; wrote the manuscript; secured the funding.

#### 443 **Funding**

444 This study was funded by a grant from the psychology department at Goldsmiths, University of London.

#### 445 **Acknowledgments**

446 We are grateful to Prof. Maria Chait and Sijia Zhao for inspiring discussions.

#### 447 **Data Availability Statement**

448 The datasets for this study can be found in the OSF repository (link:  
449 [https://osf.io/u2qdr/?view\\_only=7a9fa6bacbb249b090282377c1542d29](https://osf.io/u2qdr/?view_only=7a9fa6bacbb249b090282377c1542d29)).

#### 450 **Informed Consent Statement**

451 All participants provided written informed consent. The study was approved by the Ethics Committee  
452 at Goldsmiths, University of London.

#### 453 **References**

- 454 Alamia, A., VanRullen, R., Pasqualotto, E., Mouraux, A., & Zenon, A. (2019). Pupil-linked arousal  
455 responds to unconscious surprisal. *The Journal of Neuroscience*, 3010–18.
- 456 Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine  
457 function: Adaptive Gain and Optimal Performance. *Annual Review of Neuroscience*, 28(1), 403–  
458 450.
- 459 Barascud, N., Pearce, M., Griffiths, T., Friston, K., & Chait, M. (2016). Brain responses in humans  
460 reveal ideal-observer-like sensitivity to complex acoustic patterns. *Proceedings of the National  
461 Academy of Sciences*, 113(5), E616–25.
- 462 Bianco, R., Novembre, G., Keller, P. E., Kim, S.-G., Scharf, F., Friederici, A. D., ... Sammler, D. (2016).  
463 Neural networks for harmonic structure in music perception and action. *NeuroImage*, 142.
- 464 Bianco, R., Novembre, G., Keller, P. E., Seung-Goo, K., Scharf, F., Friederici, A. D., ... Sammler, D.  
465 (2016). Neural networks for harmonic structure in music perception and action. *NeuroImage*,  
466 142, 454–464.

- 467 Bigand, E., & Poulin-Charronnat, B. (2006). Are we “experienced listeners”? A review of the musical  
468 capacities that do not depend on formal musical training. *Cognition*, *100*(1), 100–130.
- 469 Carrus, E., Pearce, M. T., & Bhattacharya, J. (2013). Melodic pitch expectation interacts with neural  
470 responses to syntactic but not semantic violations. *Cortex*, *49*(8), 2186–200.
- 471 Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive  
472 science. *The Behavioral and Brain Sciences*, *36*(3), 181–204.
- 473 Damsma, A., & van Rijn, H. (2017). Pupillary response indexes the metrical hierarchy of unattended  
474 rhythmic violations. *Brain and Cognition*, *111*, 95–103.
- 475 Dayan, P., & Yu, A. J. (2006). Phasic norepinephrine: A neural interrupt signal for unexpected events.  
476 *Network: Computation in Neural Systems*, *17*(4), 335–350.
- 477 Dean, R. T., & Pearce, M. T. (2016). Algorithmically-generated Corpora that use Serial Compositional  
478 Principles Can Contribute to the Modeling of Sequential Pitch Structure in Non-tonal Music.  
479 *Empirical Musicology Review*, *11*(1), 27.
- 480 Eerola, T., & Toiviainen, P. (2004). MIDI Toolbox: MATLAB Tools for Music Research. University of  
481 Jyväskylä: Kopijyvä, Jyväskylä, Finland.
- 482 Fink, L. K., Hurley, B. K., Geng, J. J., & Janata, P. (2018). A linear oscillator model predicts dynamic  
483 temporal attention and pupillary entrainment to rhythmic patterns. *Journal of Eye Movement  
484 Research*, *11*(2), 12.
- 485 Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of  
486 London. Series B, Biological Sciences*, *360*(1456), 815–36.
- 487 Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., & Pantev, C. (2004). Musical training enhances automatic  
488 encoding of melodic contour and interval structure. *Journal of Cognitive Neuroscience*, *16*(6),  
489 1010–21.
- 490 Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: a review of  
491 underlying mechanisms. *Clinical Neurophysiology : Official Journal of the International  
492 Federation of Clinical Neurophysiology*, *120*(3), 453–63.
- 493 Garrido, M. I., Sahani, M., & Dolan, R. J. (2013). Outlier Responses Reflect Sensitivity to Statistical  
494 Structure in the Human Brain. *PLoS Computational Biology*, *9*(3).
- 495 Hansen, N. C., & Pearce, M. T. (2014). Predictive uncertainty in auditory sequence processing.  
496 *Frontiers in Psychology*, *5*, 1052.
- 497 Hsu, Y.-F., Le Bars, S., Hamalainen, J. A., & Waszak, F. (2015). Distinctive Representation of  
498 Mispredicted and Unpredicted Prediction Errors in Human Electroencephalography. *Journal of  
499 Neuroscience*, *35*(43), 14653–14660.
- 500 Huron, D. (2001). Tone and Voice: A Derivation of the Rules of Voice-Leading from Perceptual  
501 Principles. *Music Perception*, *19*(1), 1–64.
- 502 Huron, D. (2006). *Sweet Anticipation : Music and the Psychology of Expectation* by David Huron. (M.  
503 T. M. Press., Ed.), *Sweet Anticipation: Music and the Psychology of Expectation*. Cambridge.
- 504 Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between Pupil Diameter and  
505 Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. *Neuron*, *89*(1), 221–  
506 234.
- 507 Koelsch, S. (2016). Under the hood of statistical learning: A statistical MMN reflects the magnitude of  
508 transitional probabilities in auditory sequences. *Scientific Reports*, (August 2015), 1–11.

- 509 Koelsch, S., Gunter, T. C., v. Cramon, D. Y., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach  
510 Speaks: A Cortical “Language-Network” Serves the Processing of Music. *NeuroImage*, *17*(2),  
511 956–966.
- 512 Koelsch, S., Schmidt, B.-H., & Kansok, J. (2002a). Effects of musical expertise on the early right  
513 anterior negativity: an event-related brain potential study. *Psychophysiology*, *39*(5), 657–663.
- 514 Koelsch, S., Schmidt, B. H., & Kansok, J. (2002b). Effects of musical expertise on the early right  
515 anterior negativity: An event-related brain potential study. *Psychophysiology*, *39*(5), 657–663.
- 516 Krumhansl, C. L. (2015). Statistic, structures and style in music. *Music Perception*, *33*(1), 20–31.
- 517 Laeng, B., Eidet, L. M., Sulutvedt, U., & Panksepp, J. (2016). Music chills: The eye pupil as a mirror to  
518 music’s soul. *Consciousness and Cognition*, *44*, 161–178.
- 519 Laeng, B., Sirois, S., & Gredeback, G. (2012). Pupillometry: A Window to the Preconscious?  
520 *Perspectives on Psychological Science*, *7*(1), 18–27.
- 521 Levitin, D. J., & Tirovolas, A. K. (2009). Current advances in the cognitive neuroscience of music.  
522 *Annals of the New York Academy of Sciences*, *1156*, 211–231.
- 523 Liao, H. I., Yoneya, M., Kidani, S., Kashino, M., & Furukawa, S. (2016). Human pupillary dilation  
524 response to deviant auditory stimuli: Effects of stimulus properties and voluntary attention.  
525 *Frontiers in Neuroscience*, *10*(FEB).
- 526 Liao, H., Yoneya, M., Kashino, M., & Furukawa, S. (2018). Pupillary dilation response reflects  
527 surprising moments in music. *Journal of Eye Movement Research*, *11*(2), 1–13.
- 528 Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca’s  
529 area: an MEG study. *Nature Neuroscience*, *4*(5), 540–545.
- 530 Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of*  
531 *Neuroscience Methods*, *164*, 177–190.
- 532 Marmel, F., Tillmann, B., & Delbé, C. (2010). Priming in melody perception: tracking down the  
533 strength of cognitive expectations. *Journal of Experimental Psychology. Human Perception and*  
534 *Performance*, *36*(4), 1016–1028.
- 535 Mathôt, S., Fabius, J., Van Heusden, E., & Van der Stigchel, S. (2018). Safe and sensible preprocessing  
536 and baseline correction of pupil-size data. *Behavior Research Methods*, *50*(1), 94–106.
- 537 Meyer, L. B. (2001). Music and emotion: Distinction and uncertainties. In *Music and emotion: Theory*  
538 *and research*. (pp. 341–360). Meyer, Leonard B.: Dept of Music, U Pennsylvania, 210 SO 34th St,  
539 Philadelphia, PA, US, 19104: Oxford University Press.
- 540 Miranda, R. A., & Ullman, M. T. (2007). Double dissociation between rules and memory in music: An  
541 event-related potential study. *NeuroImage*, *38*(2), 331–345.
- 542 Müllensiefen, D., Gingras, B., Musil, J., Stewart, L., Levitin, D., Hallam, S., ... Winner, E. (2014). The  
543 Musicality of Non-Musicians: An Index for Assessing Musical Sophistication in the General  
544 Population. *PLoS ONE*, *9*(2), e89642.
- 545 Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic  
546 research of central auditory processing: A review. *Clinical Neurophysiology*, *118*(12), 2544–  
547 2590.
- 548 Narmour, E. (2015). Toward a unified theory of the I-R model (part 1): Parametric scales and their  
549 analogically isomorphic structures. *Music Perception*, *33*(1), 32–69.
- 550 Oechslin, M. S., Van De Ville, D., Lazeyras, F., Hauert, C.-A., & James, C. E. (2013). Degree of musical

- 551 expertise modulates higher order brain functioning. *Cerebral Cortex (New York, N.Y. : 1991)*,  
552 23(9), 2213–24.
- 553 Omigie, D., Pearce, M., Lehongre, K., Hasboun, D., Navarro, V., Adam, C., & Samson, S. (2019).  
554 Intracranial Recordings and Computational Modeling of Music Reveal the Time Course of  
555 Prediction Error Signaling in Frontal and Temporal Cortices. *Journal of Cognitive Neuroscience*,  
556 31(6), 855–873.
- 557 Omigie, D., Pearce, M. T., & Stewart, L. (2012). Tracking of pitch probabilities in congenital amusia.  
558 *Neuropsychologia*, 50(7), 1483–1493.
- 559 Omigie, D., Pearce, M. T., Williamson, V. J., & Stewart, L. (2013). Electrophysiological correlates of  
560 melodic processing in congenital amusia. *Neuropsychologia*, 51(9), 1749–1762.
- 561 Pearce, M.T. (2005). The Construction and Evaluation of Statistical Models of Melodic Structure in  
562 Music Perception and Composition. *Dissertation*, (December), 267.
- 563 Pearce, M.T. (2018). Statistical learning and probabilistic prediction in music cognition: Mechanisms  
564 of stylistic enculturation. *Annals of the New York Academy of Sciences*, 1423, 378–395.
- 565 Pearce, M.T., Ruiz, M. H., Kapasi, S., Wiggins, G. a, & Bhattacharya, J. (2010). Unsupervised statistical  
566 learning underpins computational, behavioural, and neural manifestations of musical  
567 expectation. *NeuroImage*, 50(1), 302–13.
- 568 Pearce, Marcus T, & Wiggins, G. a. (2006). Expectation in melody: the influence of context and  
569 learning. *Music Perception*, 23(45), 377–405.
- 570 Pearson, R. K., Neuvo, Y., Astola, J., & Gabbouj, M. (2016). Generalized Hampel Filters. *Eurasip*  
571 *Journal on Advances in Signal Processing*, 2016(1).
- 572 Quiroga-martinez, D. R., Hansen, N. C., Højlund, A., Pearce, M., Brattico, E., Vuust, P., & Quiroga-  
573 martinez, D. R. (2019). Decomposing neural responses to melodic surprise in musicians and  
574 non-musicians : evidence for a hierarchy of predictions in the auditory system, 1–26.
- 575 Ricardo Quiroga-Martinez, D. (2018). Precision-weighting of musical prediction error: Converging  
576 neurophysiological and behavioral evidence, 1–24.
- 577 Rohrmeier, M., Rebuschat, P., & Cross, I. (2011). Incidental and online learning of melodic structure.  
578 *Consciousness and Cognition*, 20(2), 214–222.
- 579 Ross, S., & Hansen, N. C. (2016). Dissociating Prediction Failure: Considerations from Music  
580 Perception. *Journal of Neuroscience*, 36(11), 3103–3105.
- 581 Rubin, J., Ulanovsky, N., Nelken, I., & Tishby, N. (2016). The Representation of Prediction Error in  
582 Auditory Cortex. *PLOS Computational Biology*, 12(8), e1005058.
- 583 Sara, S. J. (2009). The locus coeruleus and noradrenergic modulation of cognition. *Nature Reviews*.  
584 *Neuroscience*, 10(3), 211–223.
- 585 Sohoglu, E., & Chait, M. (2016). Detecting and representing predictable structure during auditory  
586 scene analysis. *ELife*, 5(Se), 1–17.
- 587 Southwell, R., Baumann, A., Gal, C., Barascud, N., Friston, K., & Chait, M. (2017). Is predictability  
588 salient? A study of attentional capture by auditory patterns. *Philosophical Transactions of the*  
589 *Royal Society B: Biological Sciences*, 372(1714).
- 590 Southwell, R., & Chait, M. (2018). Enhanced deviant responses in patterned relative to random sound  
591 sequences. *Cortex*, 109, 92–103.
- 592 Tervaniemi, M., Tupala, T., & Brattico, E. (2012). Expertise in folk music alters the brain processing of

593 Western harmony. *Annals of the New York Academy of Sciences*, 1252(1), 147–151.

594 Tervaniemi, Mari. (2009). Musicians-Same or Different? *Annals of the New York Academy of Sciences*,  
595 1169(1), 151–156.

596 Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing  
597 approach. *Psychological Review*, 107(4), 885–913.

598 Tillmann, B., Janata, P., & Bharucha, J. J. (2003). Activation of the inferior frontal cortex in musical  
599 priming. *Cognitive Brain Research*, 16(2), 145–161.

600 Tillmann, B., & Lebrun-Guillaud, G. (2006). Influence of tonal and temporal expectations on chord  
601 processing and on completion judgments of chord sequences. *Psychological Research*, 70(5),  
602 345–58.

603 Troncoso, X. G., Macknik, S. L., & Martinez-Conde, S. (2008). Microsaccades counteract perceptual  
604 filling-in. *Journal of Vision*, 8(14), 15–15.

605 Vuust, P., Brattico, E., Seppänen, M., Näätänen, R., & Tervaniemi, M. (2012a). Practiced musical style  
606 shapes auditory skills. *Annals of the New York Academy of Sciences*, 1252(1), 139–146.

607 Vuust, P., Brattico, E., Seppänen, M., Näätänen, R., & Tervaniemi, M. (2012b). The sound of music:  
608 Differentiating musicians using a fast, musical multi-feature mismatch negativity paradigm.  
609 *Neuropsychologia*, 50(7), 1432–1443.

610 Vuust, P., & Witek, M. a. G. (2014). Rhythmic complexity and predictive coding: a novel approach to  
611 modeling rhythm and meter perception in music. *Frontiers in Psychology*, 5, 1111.

612 Weiskrantz, L., Cowey, A., & Barbur, J. L. (1999). Differential pupillary constriction and awareness in  
613 the absence of striate cortex. *Brain*, 122(8), 1533–1538.

614 Wetzel, N., Buttelmann, D., Schieler, A., & Widmann, A. (2016). Infant and adult pupil dilation in  
615 response to unexpected sounds. *Developmental Psychobiology*, 58(3), 382–392.

616 Widmann, A., Schröger, E., & Wetzel, N. (2018). Emotion lies in the eye of the listener: Emotional  
617 arousal to novel sounds is reflected in the sympathetic contribution to the pupil dilation  
618 response and the P3. *Biological Psychology*, 133(January), 10–17.

619 Xing, B., Li, Y. C., & Gao, W. J. (2016). Norepinephrine versus dopamine and their interaction in  
620 modulating synaptic function in the prefrontal cortex. *Brain Research*, 1641, 217–233.

621 Zatorre, R. J., & Salimpoor, V. N. (2013). From perception to pleasure: music and its neural  
622 substrates. *Proceedings of the National Academy of Sciences*, 110, 10430–7.

623 Zhao, S., Chait, M., Dick, F., Dayan, P., Furukawa, S., & Liao, H.-I. (2019). Pupil-linked phasic arousal  
624 evoked by violation but not emergence of regularity within rapid sound sequences. *Nature*  
625 *Communications*, 10, 4030.

626

627