

Incremental comprehension of pitch relationships in written music:

Evidence from eye movements

Running Head: Incremental comprehension of music

Lauren V. Hadley, Patrick Sturt, Tuomas Eerola, and Martin J. Pickering

Lauren V. Hadley, Department of Psychology, University of Edinburgh; Patrick Sturt, Department of Psychology, University of Edinburgh; Tuomas Eerola, Department of Music, Durham University; Martin J. Pickering, Department of Psychology, University of Edinburgh.

Address for correspondence:

Lauren Victoria Hadley

Department of Psychology

University of Edinburgh

Edinburgh EH8 9JZ

UK

E-mail: lauren.hadley@cantab.net

phone: +44 (0) 7903 219 015

fax: +44 (0) 131 650 3461

Acknowledgements: L. V. Hadley was supported by a grant from the Economic and Social Research Council. We would like to thank Dr Mateo Obregón for invaluable help with data analysis, and Marcus Johnson for extensive assistance with eye tracking matters.

Abstract

To investigate how proficient pianists comprehend pitch relationships in written music when they first encounter it we conducted two experiments in which proficient pianists' eyes were tracked while they read and played single-line melodies. In Experiment 1, participants played at their own speed; in Experiment 2 they played with an external metronome. The melodies were either congruent or anomalous, with the anomaly involving one bar being shifted in pitch to alter the implied harmonic structure (e.g., non-resolution of a dominant). In both experiments, anomaly led to rapid disruption in participants' eye-movements in terms of regressions from the target bar, indicating that pianists process written pitch relationships online. This is particularly striking because in musical sight-reading eye movement behaviour is constrained by the concurrent performance. Both experiments also showed that anomaly induced pupil dilation. Together these results indicate that proficient pianists rapidly integrate the music that they read into the prior context, and that anomalies in terms of pitch relationships lead to processing difficulty. These findings parallel those of text reading, suggesting that structural processing involves similar constraints across domains.

Keywords: Sight-reading; Music performance; Music processing; Eye movements

Incremental comprehension of written music: Evidence from eye movements

Listeners process heard music incrementally, relating each new musical event to the prior context (Janata, 1995; Koelsch et al., 2001). But do musicians playing from a score comprehend such musical relationships as they read, and therefore experience disruption if those relationships are anomalous? In this paper, we report two sight-reading experiments that introduce pitch-based anomalies into melodies and investigate whether such anomalies lead to rapid disruptions in pianists' eye movements, just as anomalous words lead to rapid disruption in text readers' eye movements (Rayner, 1998).

If written music is processed similarly to written language, we would expect notation to be processed incrementally during sight-reading. Much evidence for incremental processing of text comes from studies investigating readers' eye movements. Eye movements index cognitive processing, with responses to visually encountering a word reflecting the ease with which that word is processed (Rayner & Duffy, 1986). Initial processing difficulty is commonly indicated by a high proportion of backward glances (regressions) from the anomalous region and long reading times in the region.

For example, Just and Carpenter (1980) found that readers' eye movements were immediately disrupted by texts that were difficult to interpret. More recently, Rayner et al. (2004) had participants read sentences such as:

- 1a. Control:* John used a knife to chop the large carrots for dinner
- 1b. Mild Implausibility:* John used an axe to chop the large carrots for dinner
- 1c. Severe Implausibility:* John used a pump to inflate the large carrots for dinner

In (1b-c), the implausibility becomes apparent on the word *carrots* (the "target"). Participants spent longer fixating *carrots* when they first encountered it in (1c) than (1a), indicating that

they rapidly integrated it with the prior context. The mildly implausible condition (1b) led to more regressions from the following words *for dinner* (the “post-target region”) than (1a). In reading, then, integration difficulty may lead to increases in processing time and regressions, which occur later with mild than severe anomalies (see also Rayner, Chace, Slattery, & Ashby, 2006; Warren & McConnell, 2007). Similar effects occur during the reading of syntactic anomalies (Frazier & Rayner, 1982).

We suggest that if musical relationships are interpreted like semantic and syntactic relationships in language, we would find similar rapid effects of anomaly. In standard tonal Western music, an analogous anomaly would be a ‘wrong note’ or ‘inappropriate progression’. In such music, the relationships between concurrent or successive pitches form norms against which anomalous notes and progressions can be judged. These norms (based on groups of notes) form the basis of harmonic syntax, which many researchers have treated as equivalent to linguistic contextual constraint (see Patel, 2003).

In this study we focus on sight-reading: the unprepared performance of music from a score (Hallam, Cross, & Thaut, 2009). We investigate reading with concurrent performance because it is hard to be certain how musicians process notation when they are not playing (e.g. in terms of visual properties or to detect passages of difficulty). When sight-reading, however, the aim is to play each note at the right time, and this presumably requires interpretation of the notation in the order that it is written (similarly to text reading).

Sight-reading studies have primarily examined coarse-grain eye movement measures such as the perceptual span and eye-hand span (e.g. Gilman & Underwood, 2003; Truitt et al., 1997; Furneaux & Land, 1999; Rosemann, Altenmuller & Fahle, 2016). Other work has focused on how notational complexity affects eye movement during music reading in novices compared to experts. Goolsby (1994a, b) found that pianists make shorter fixations and

saccades when reading notation with a larger number of notes and performance markings, compared to notation that is sparser. Wurtz et al. (2009) showed that violinists make shorter fixations for rhythmically regular pieces than rhythmically variable pieces. However, as the rhythmically regular piece was more visually homogenous and included fewer performance markings, these differences may again be due to differences in the visual properties of the notation. In fact, there has been little empirically controlled examination of how musical content (as opposed to visual features of notation) is processed using eye tracking (Madell & Hebert, 2008). In this study we investigate how proficient pianists interpret musical relationships when first reading a novel melody.

One study of music reading focused on the processing of a memorized melody (Penttinen, Huovinen, & Ylitalo, 2014). In their experiment, the notation for ‘Mary had a little lamb’ was presented to pianists in either its original version, or an anomalous version involving one bar being shifted down a tone. At the anomalous bar, participants showed a decreased eye-hand span, which, despite being an indirect measure of processing difficulty, suggests that the expectancy violation in a familiar melody causes disruption.

Although Penttinen et al’s (2014) study shows that pitch anomalies can cause processing difficulty, the design involved participants processing anomalies between the notation of a tune, and their auditory memory of that tune. Their findings may therefore be due to the mismatch between the notation and participants’ long-term representation of the tune, rather than the comprehension of a contextually anomalous pitch relationship. We were interested in how pitch relationships are interpreted when a novel piece is read for the first time, to understand whether such relationships are comprehended incrementally or after a delay.

A previous study investigating the effect of anomalies in novel pieces during music reading tracked pianists' eye movements while they read musical notation and sentence stimuli (Ahken, Comeau, Hébert, & Balasubramaniam, 2012). Half of the musical stimuli were presented with key signatures and half with accidentals. Whereas key signatures define the tonally appropriate note-set at the start of the piece, the use of accidentals means that the C-major note-set is used by default, and notes external to this set are created using transient symbols called accidentals. In sentence reading, participants' total fixation time was longer on anomalous than congruent target words (in a two-word region). They showed a similar effect of anomaly for the musical stimuli written with key signatures, but not for the stimuli written with accidentals.

There are several methodological concerns with the study. In the key signature condition, only the anomalous stimuli involved additional symbols (accidentals), thereby confounding visual complexity with musical anomaly. Additionally, there was no control for participants' timing, and hence any differences in eye movements may have been related to differences in performance (i.e., people may have taken longer to read anomalous stimuli because they were playing more slowly, rather than because of difficulty processing the anomaly). Moreover, they reported only late processing measures (i.e., measures that included re-reading time after participants had potentially read the whole stimulus). The authors therefore provided no evidence that anomalies are processed differently to congruent melodies during initial processing.

In sum, neither of the previous studies investigating pitch anomalies during music reading (Ahken et al., 2012; Penttinen et al., 2014) indicate whether contextual pitch relationships are processed online when reading a novel melody. Furthermore, one study used a pattern-matching paradigm (involving sequential presentation of two excerpts of notation that were either the same or different) to investigate the processing of simple versus complex

pitch relationships over repeated presentations (Waters & Underwood, 1998). This study found no complexity effect when the stimuli were first encountered (i.e., in presentation of the first stimulus), though there was an effect of complexity when the comparison stimulus was presented.

Our study therefore investigated how musical pitch relationships are processed on-line as they are first encountered, in order to infer the level of comprehension when musicians first encounter a piece. We used contextual musical anomalies in novel musical melodies to investigate comprehension during sight-reading, and focused on the effects of expectation generated through a single-line melody. We were particularly interested in determining whether pitch relationships in written music are incrementally processed. If music is processed incrementally, in a manner similar to language, we would expect effects of anomaly to occur soon after it is encountered. But it is also possible that the constraints on reading rate caused by concurrent performance lead musicians to read written notes simply as instructions for action. If so, we predict that the processing of pitch relationships will be delayed and effects of anomaly will be delayed

In our experiments we based our congruence manipulation on changing the pitch to alter the implied harmonic structure underlying the melodies. We applied this manipulation at the level of the musical bar, and also used the musical bar as our level of analysis, because expert musicians use a chunking strategy to process groups of notes at a time especially when they outline chords (Waters, Underwood, & Findlay, 1997; Polanka, 1995). Moreover, the bar-level has previously been used for eye movement analyses (Drai-Zerbib, Baccino, & Bigand, 2011; Drai-Zerbib & Baccino, 2013). Stimuli were constructed to be of equivalent notational complexity in the congruent and anomalous conditions. In Experiment 1 pianists played the melodies at their chosen tempo to investigate natural reading behaviour, and in

Experiment 2 pianists played in time with a metronome to ensure that effects in Experiment 1 were not dependent on expressive performance choices.

We hypothesise that processing difficulty will be evident rapidly after reading a musical anomaly. Studies of language processing have shown such difficulty to become evident through the rate of regression and duration of initial reading time, in or immediately after a short target region (see Rayner et al., 2004; Stewart et al., 2004). We therefore predict that anomalous stimuli will show an increased rate of regression and longer initial reading times than the non-anomalous stimuli in and immediately after the target bar. As music reading is guided by the temporal framework of the piece being played, and delays in moving the eyes would lead to performance errors, we ran an additional exploratory analysis of pupil dilation. This is a neurophysiological index of cognitive effort that has been used in many fields including language processing (Schluroff, 1982; Sevilla, Maldonado, & Shalóm, 2014), and may be particularly relevant for music reading since reading rate is relatively constrained by performance speed (Rayner & Pollatsek, 1997; Truitt, Clifton, Pollatsek, & Rayner, 1997). We predict that anomalous melodies may lead to greater pupil dilation than congruent melodies after the point of anomaly (see Schluroff, 1982; Sevilla et al., 2014). Finally, we analysed performance errors in terms of both timing and pitch, to investigate whether any eye movement effects were dependent on performance differences across conditions.

Experiment 1 (no metronome)

Method

Participants

Thirty active pianists took part in the study. However three were removed because of recording issues, and three according to the exclusion criteria stated below. Of the remaining 24, 17 were female, with an age range from 18 to 66 ($M=43.25$, $SD=17.09$). A standard

musical questionnaire was used to record pianists' musical training (Goldsmiths Musical Sophistication Index; Müllensiefen, Gingras, Stewart, & Musil, 2011). All participants had nine or more years of formal musical tuition. Furthermore, 20 had undertaken instrumental tuition for over 10 years, three for six to nine years, and one for two years (years of instrumental training was not correlated with age; $r(24) = .186, p = .383$).

Materials

The experiment involved 16 items, each being composed in four versions (see Figure 1). For each item, we constructed two melodic contexts, and two versions of the target bar (shifted to start on different notes). The targets at each pitch were related through transposition, and hence matched in contour. Within each item, both versions of the target bar were used in a congruent and anomalous version, and every melodic context was used in both a congruent and anomalous version (thus avoiding any differences in the stimuli between conditions). Figure 1 presents an example item used in the experiment. In the congruent stimuli, the pitches in the target bar formed a probable progression from the previous bar and hence congruent harmonic structure. In the anomalous stimuli, the pitches in the target bar formed an improbable progression from the previous bar and hence anomalous harmonic structure, but remained appropriately within the key of the piece (i.e., they were unlikely but still possible according to harmony norms). Such constraints meant that the target bar more frequently followed the previous bar by step in the congruous stimuli (i.e., moving between consecutive notes), and by leap in the anomalous stimuli (i.e., moving between non-consecutive notes, though the interval was within a third in all but 5 anomalous stimuli).

(Figure 1 about here)

The first author (a trained pianist and music graduate) composed a candidate set of 25 items. All were 8 bars in length to be played in the right hand of the piano. Expectation for

the congruent target was set up through use of an implied dominant chord in the pre-target bar that required resolution. To select the items, 22 instrumentalists rated sets of melodies according to their naturalness on a 1 (very unnatural) to 7 (very natural) scale. We selected the 16 items in which the congruent and anomalous versions differed as much as possible. These differences ranged from 0.36 to 1.5 ($M=0.93$, congruent $M=4.59$, anomalous $M=3.80$). They were relatively small because of the need to avoid accidentals in the target bar (which would have led to a visual confound between conditions as in Ahken et al., 2012). The target was positioned in the 4th, 5th, 6th, or 7th bar. Within an item, the largest transpositional difference between target bars was a perfect fifth. We additionally constructed 32 eight-bar fillers in a similar style. Items were placed in two lists of 32 stimuli, each containing one congruent and one anomalous version of each item (1 & 3, or 2 & 4), to allow a within-participant analysis of the effect of congruency. These stimuli comprised 10 melodies in 3_4 time, 10 in 4_4 time, and 12 in 6_8 time. The experiment was presented in two halves, with one version of each item in each half. Half of the participants were randomly assigned to each list.

Procedure

Participants were seated in front of a midi keyboard set and eye movements were recorded using a SR EyeLink 1000 remote-mode eye-tracker, sampling the left eye at 500Hz. Participants were asked to begin playing as soon as possible after presentation of the notation (as opposed to skimming the music before playing). They were told to continue without correction if they made a mistake. If participants asked the speed at which to play, they were told to choose a tempo that felt natural (no minimum or maximum was specified). The experiment lasted approximately an hour.

Data Analysis

Participants' performances were analysed by the primary author for pitch errors. As pitch errors may have led to stimuli being processed as anomalous even if congruent, or vice

versa, we removed any participants that made pitch errors in at least 50% experimental trials (2 participants). We also removed any participants that fixated beyond the third bar of music before 2 seconds had elapsed in at least 25% of experimental trials, as this suggested pre-processing of the stimuli (1 participant). The remaining participants made one or more errors on a mean of 7.08 experimental trials (i.e., 22% of trials), with 4.92 experimental trials specifically incorporating a pitch error (i.e., 15% of trials). Participants were equally likely to make errors on congruent and anomalous trials ($t_1(23) = -0.63, p = .53$). [We estimated timing errors by comparing the onsets of the performed pitches to the underlying score by matching the two with a dynamic time warping algorithm \(see Large & Rankin, 2007\).](#) Overall, the tempo the participants chose was moderate ($M=93.90$ beats per minute, $SD=25.5$).

Raw eye-position data were parsed using psychophysical parsing thresholds adopted by the EyeLink DataViewer™ software (SR Research Ltd.). According to these thresholds, we defined a fixation as a series of samples in which the eye did not move more than 0.2 degrees, with a velocity of under 30 deg/s. [Time during blinks did not contribute to fixation or pupil measures.](#)

Results

The stimuli were divided into three regions of interest: (a) pretarget bar, (b) target bar, and (c) post-target bar (see Figure 2).

(Figure 2 about here)

We analysed the data using a standard set of eye-tracking measures (see Rayner, 1998). These included *first-pass reading time*: summed fixation times from the beginning of the first fixation in a region to the end of the last fixation in the region before the eye leaves that region (i.e., excluding saccade times); *first-pass regressions out of a region*: proportion

of trials in which the region is exited to the left prior to leaving the region to the right; and *total reading time*: summed times of all fixations in a region (i.e., including re-reading after exiting the region). In addition, we computed *average pupil size*: pupil dilation averaged over all fixations on a region (i.e., over total time). Pupil size as reported by the EyeLink 1000 is a calculation of area based on the number of thresholded pixels in the pupil image (pixel area or *pa*), and is a ratio scale that relates to pupil diameter in mm (Hayes & Petrov, 2016). We note that pupil size recordings are affected by participant setup and eye position (Pomplun & Sindhura, 2003), meaning that a participant's movement may show as dilation changes, and hence specify that these analyses are exploratory. For early processing measures (first-pass reading time and first-pass regressions out), regions not fixated on first pass were excluded, following standard practice. For reading-time and pupil-dilation measures, any data 3 standard deviations from each individual participant's mean were excluded from analysis.

We expected no difference between conditions in the pretarget region, as the target bar had not yet been processed. In the target and post-target bars we expected early disruption in the anomalous condition as demonstrated by a longer first-pass reading time, and a higher rate of first-pass regression. In the additional exploratory investigation of pupilometry we expected greater pupil dilation following the anomalous than the congruent target.

Paired t-tests were computed between the congruent and anomalous conditions by participants (t_1) and items (t_2). The target bar was fixated during first-pass reading of 96% trials (95% in congruent and 96% in anomalous trials), and during first-pass it elicited a mean of 2.64 fixations (2.63 in congruent and 2.64 in anomalous trials). This was a mean of one fixation every 1.66 notes, similar to previous research (Rayner & Pollatsek, 1997). This finding supported the use of a bar- rather than note-based analysis, as it meant that most regions-of-analysis contained one or more fixations. Moreover, there were no significant

correlations in the target bar between number of fixations on first pass and number of notes ($r(16) = -.341, p = .197$), or number of beats ($r(16) = .116, p = .668$), suggesting that number of fixations in a bar is relatively independent of its content. We additionally found no difference in the amount of time spent on individual trials between conditions (congruent $M = 20.86s$, anomalous $M = 21.04s$; $t_1(23) = -1.62, p = .12$; $t_2(15) = -1.21, p = .25$), suggesting that condition did not affect overall reading speed. See Table 1.

(Table 1 about here)

As predicted, there were no significant differences between conditions in the pretarget bar ($ps > .05$), but participants were more likely to regress out of the target bar in the anomalous than the congruent condition ($M \text{ diff} = .07$; $t_1(23) = -2.27, p = .03$; $t_2(15) = -2.13, p = .05$). Furthermore, as predicted in our exploratory analysis of pupil dilation, in the post-target bar mean pupil size was significantly greater in the anomalous than the congruent condition ($M \text{ diff} = 4pa$; $t_1(23) = -2.12, p = .05$; $t_2(15) = -2.17, p = .05$). No other effects were significant. There were no differences in performance of the congruent and anomalous melodies in terms of timing error or pitch error in any region (all $ps \geq .15$). See Table 1.

Discussion

The increased rate of regression from the target bar in the anomalous compared to congruent condition indicates that anomaly disrupted processing very rapidly. The greater pupil dilation in the anomalous post-target bar is a delayed response that also relates to processing difficulty. These effects occurred in the absence of any difference in performance between the conditions. Overall, it therefore appears that proficient pianists incrementally interpret written music, with processing of contextual relationships occurring in spite of the requirement to read in a temporally constrained manner (due to concurrent performance). Our second experiment considered whether effects of anomaly would remain when reading rate

was explicitly constrained by a metronome, both controlling performance timing and increasing cognitive demands.

Experiment 2 (with metronome)

Experiment 2 was identical to Experiment 1, except that a metronome was used.

Participants

Thirty-three further pianists took part in the study. However, nine were removed due to the exclusion criteria. Of the remaining 24, half were female, with an age range from 18 to 69 ($M=34.08$, $SD=14.23$). All participants had taken six or more years of formal musical tuition. Furthermore, regarding instrumental training, 20 had undertaken tuition for over 10 years, and four for six to nine years (uncorrelated with age; $r(24) = .091$, $p = .673$).

Participants remained active pianists at the time of the experiment.

Method

Materials

These were the same as in Experiment 1.

Procedure

The procedure was the same as in Experiment 1, except that participants were asked to play in time with a metronome that clicked on each beat (in all time signatures an eighth note was timed to be 300ms; but in $\frac{3}{4}$ and $\frac{4}{4}$ stimuli a click occurred every quarter note, in $\frac{6}{8}$ stimuli a click occurred every half note). We judged this to be a moderate tempo requiring concentration yet closely matching the tempo spontaneously used in Exp. 1 (mean inter-beat-interval of 318ms). Participants were asked to begin playing as soon as possible after the notation appeared and the concurrent clicks began (i.e., there were no ‘cue’ clicks), and to keep to the beat of the metronome.

Data Analysis

The criteria for removal of data were the same as in Experiment 1. Of the nine participants whose data were removed, two were because of performance errors (suggesting lack of sight-reading skill) and seven because of melodic preview (suggesting pre-processing). The remaining participants made one or more errors on a mean of 7.25 experimental trials (i.e., 23% of trials), with 6.13 experimental trials specifically incorporating a pitch error (i.e., 19% of trials). Participants were equally likely to make errors on congruent and anomalous trials ($t_1(23) = -0.33, p = .74$).

Results

Data were preprocessed as in Experiment 1. Target bars were fixated during first-pass reading on 95% of trials (96% in congruent and 94% in anomalous trials), and during first-pass they elicited a mean of 2.30 fixations (2.31 in congruent and 2.30 in anomalous trials). This was a mean of one fixation every 1.90 notes, and in the target bar we again found no significant correlations between number of fixations on first pass and number of notes ($r(16) = -.310, p = .243$) or number of beats ($r(16) = .310, p = .242$). We found no difference in the amount of time spent on individual trials between conditions (congruent $M = 21.67s$, anomalous $M = 21.65s$; $t_1(23) = 0.22, p = .83$; $t_2(15) = 0.37, p = .72$), suggesting that condition did not affect overall reading speed. See Table 2.

(Table 2 about here)

As predicted, there were no significant differences between conditions in the pretarget bar ($ps > .05$), but participants were more likely to regress out of the target bar in the anomalous than the congruent condition ($M \text{ diff} = .10$; $t_1(23) = -3.97, p < .001$; $t_2(15) = -3.89, p = .001$). Furthermore, as predicted in our exploratory analysis of pupil dilation, mean pupil

size was significantly greater in the anomalous than the congruent condition in the target and post-target bars (Target bar: M diff = 7 pa; $t_1(23) = -2.18, p = .04$; $t_2(15) = -2.54, p = .02$; Post-target bar: M diff = 11 pa; $t_1(23) = -4.22, p < .001$; $t_2(15) = -3.29, p = .01$).

Two further analyses showed total times to be longer in the anomalous than the congruent target bar (M diff = 1.36s; $t_1(23) = -2.17, p = .04$; $t_2(15) = 2.12, p = .05$) and first-pass times to be longer in the congruent than the anomalous post-target bar (M diff = 1.61s; $t_1(23) = 2.18, p = .04$; $t_2(15) = 2.50, p = .03$). No other effects were significant. In addition, there was no difference in performance of the congruent and anomalous melodies in any region (all $ps \geq .15$). See Table 2.

Discussion

In the target bar, differences in the rate of regressions out across conditions showed that participants encountered rapid difficulty processing anomalous compared to congruent melodies. Exploratory pupillometry analysis also indicated increased cognitive load in the anomalous condition for both the target and the post-target bars. These effects were also found in Experiment 1, and again occurred in the absence of any difference in performance between conditions.

We additionally found a reversed effect of longer first-pass time in the congruous than anomalous post-target bar. This effect may reflect compensatory processes: when participants were disrupted by an anomaly, they had less time to process the next part of the score if they wished to keep in time with the metronome, as was their task. Increased total time in the anomalous target bar indicates later processing difficulty.

General Discussion

In two experiments, we showed that the eye-movements of proficient pianists were rapidly disrupted when reading a pitch relationship anomaly in a score (although performance timing was unaffected). As predicted, this effect was found in regression measures immediately following the anomalies. This finding indicates that sight-reading a novel melody involves incremental processing, with readers rapidly constructing a representation of the music and integrating new input into this interpretation. We did not, however, find the predicted effect of congruence on first-pass time, possibly due to the constraints of reading in time with temporally regulated performance. Finally, as predicted, exploratory analyses of pupil dilation in both experiments found that reading an anomalous musical melody led to an increase in pupil size.

We interpret our pupilometry data with caution, as such measurements are affected by the angle of the eye (Pomplun & Sindhura, 2003), light intensity (Ellis, 1981), and individual differences (e.g., Ellermeier & Westphal, 1995). Such factors are difficult to control in an ecologically valid study of music reading (in which musicians are free to move appropriately for performance). These factors were therefore not explicitly controlled in this experiment, although targets occurred in the same position within items, lighting was kept constant, and congruence was analysed within-subjects. We found that in Experiment 1 (without metronome) anomalous melodies elicited a significant increase in pupil dilation in the post-target bar. In Experiment 2 (with metronome), the effects of pupil dilation replicated in the post-target bar and also occurred in the target bar. This difference between experiments may relate to the increased cognitive demands of playing with an external timekeeper. However, we note that further investigation of the pupilometry effects should be undertaken in a constrained setup with a fixed head position.

The speed with which anomaly affects eye movements during sight-reading indicates that musicians incrementally process pitch relationships in the music that they read.¹ However, our use of sight-reading meant that participants received both motor and audio feedback of their performance. The availability of such feedback means that there are two ways in which such incremental processing could occur. Either the musician could be interpreting the pitch relationships from the score alone, or they could be interpreting pitch relationships after they hear the output of their performance (or have experienced the act of production). Previous studies suggest that the eye-hand span is approximately one beat (Gilman & Underwood, 2003; Truitt et al., 1997), and so our musicians would likely have been reading the target bar and hearing their performance of it with only a beat's delay. Our eye movement disruption could therefore have been a response to the written music or the auditory output. We suggest that as our earliest measure of processing disruption was so rapid (regressions occurring in the target bar), and did not carry over to the post-target bar (which might be expected if the processing was one beat behind), musicians were incrementally processing the written music.

It is also unclear whether musicians were incrementally interpreting the pitch of the melodies in terms of harmonic structure (underlying norms), or in terms of note-to-note progressions. We have assumed that the rapid difficulty that we have found is due to musicians interpreting implied harmonic structure, in much the same way as semantic or syntactic processing in language. It is possible that our results instead reflect incremental

¹ The effect of congruence occurring in the target bar onwards is particularly striking given the relatively small differences in congruence ratings between conditions. It is possible that this difference is comparable to Rayner et al.'s (2004) difference between control and mildly implausible conditions, which led to delayed eye movement disruption.

processing at the note-to-note level, and the fact the congruent melodies are more likely to progress by step than anomalous melodies. However, the majority of anomalous stimuli involved a leap of no larger than a third (27/32 stimuli), and a similar claim about the importance of transitional probability between words in language reading has been shown to be small and better explained by predictability (McDonald & Shillcock, 2003; Frisson, Rayner, & Pickering, 2005). Finally, a third consideration is whether differences in ease of performance may have driven the processing difficulty of the anomalous stimuli. While we controlled the content and visual information in the target bar across conditions, shifting the target bar with respect to the context may also have affected how easy the melody was to play in terms of fingering.

Our experiments show that pitch relationships in written music are interpreted incrementally by proficient pianists even when reading speed is constrained. These findings strongly suggest that during initial processing musicians comprehend notation in terms of contextual musical relationships, as opposed to simple performance instructions. The measures in which we found processing difficulty (first-pass regressions, pupil size), and the speed with which individuals responded to the anomalies (target and post-target region) show many similarities with previous research into language and linguistic anomalies. Furthermore, these effects occurred in the absence of any differences in performance timing. We therefore propose that musicians experience similar processing difficulty when reading pitch-based anomalies in music to that encountered by text readers encountering anomalies in language.

- Ahken, S., Comeau, G., Hébert, S., & Balasubramaniam, R. (2012). Eye movement patterns during the processing of musical and linguistic syntactic incongruities. *Psychomusicology: Music, Mind and Brain*, 22(1), 18–25.
- Drai-Zerbib, V., & Baccino, T. (2013). The effect of expertise in music reading: cross-modal competence. *Journal of Eye Movement Research*, 6(5), 1–10.
- Drai-Zerbib, V., Baccino, T., & Bigand, E. (2011). Sight-reading expertise: Cross-modality integration investigated using eye tracking. *Psychology of Music*, 40(2), 216–235.
- Ellermeier, W., & Westphal, W. (1995). Gender differences in pain ratings and pupil reactions to painful pressure stimuli. *Pain*, 61(3), 435-439.
- Ellis, C. J. (1981). The pupillary light reflex in normal subjects. *British Journal of Ophthalmology*, 65(11), 754-759.
- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14(2), 178–210.
- Frisson, S., Rayner, K., & Pickering, M. J. (2005). Effects of contextual predictability and transitional probability on eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(5), 862–877.
- Furneaux, S., & Land, M. F. (1999). The effects of skill on the eye–hand span during musical sight–reading. *Proceedings of the Royal Society of London B: Biological Sciences*, 266(1436), 2435-2440.

- Gilman, E., & Underwood, G. (2003). Restricting the field of view to investigate the perceptual spans of pianists. *Visual Cognition*, *10*(2), 201–232.
- Goolsby, T. W. (1994a). Eye movement in music reading: Effects of reading ability, notational complexity, and encounters. *Music Perception: An Interdisciplinary Journal*, *12*(1), 77-96.
- Goolsby, T. W. (1994b). Profiles of processing: Eye movements during sightreading. *Music Perception: An Interdisciplinary Journal*, *12*(1), 97-123.
- Hayes, T. R., & Petrov, A. A. (2016). Mapping and correcting the influence of gaze position on pupil size measurements. *Behavior Research Methods*, *48*(2), 510-527.
- Hallam, S., Cross, I., & Thaut, M. (2009). *Oxford handbook of music psychology* (pp. 344–351). Oxford, UK: Oxford University Press.
- Janata, P. (1995). ERP measures assay the degree of expectancy violation of harmonic contexts in music. *Journal of Cognitive Neuroscience*, *7*(2), 153–164.
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*(4), 329–354.
- Koelsch, S., Gunter, T. C., Schröger, E., Tervaniemi, M., Sammler, D., & Friederici, a D. (2001). Differentiating ERAN and MMN: an ERP study. *Neuroreport*, *12*(7), 1385–9.
- Large E., & Rankin, S. (2007). Matching a Performance to Notation. In T. Eerola and P. Toiviainen (Eds.), *MIDI Toolbox: MATLAB Tools for Music Research*, University of Jyväskylä, Kopijyvä, Jyväskylä, Finland.

- Madell, J., & Hébert, S. (2008). Eye movements and music reading: Where do we look next? *Music Perception: An Interdisciplinary Journal*, 26(2), 157-170.
- McDonald, S. A., & Shillcock, R. C. (2003). Eye movements reveal the on-line computation of lexical probabilities during reading. *Psychological Science*, 14(6), 648–652.
- Müllensiefen, D., Gingras, B., Stewart, L., & Musil, J. (2011). The Goldsmiths Musical Sophistication Index (Gold-MSI): Technical Report and Documentation v0. 9. *London: Goldsmiths, University of London*.
- Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, 6(7), 674–81.
- Penttinen, M., Huovinen, E., & Ylitalo, A. (2014). Reading ahead: adult music students' eye movements in temporally controlled performances of a children's song. *Journal of Music Education*, 33(1), 36–50.
- Polanka, M. (1995). Research note: Factors affecting eye movements during the reading of short melodies. *Psychology of Music*, 23(2), 177-183.
- Pomplun, M., & Sunkara, S. (2003). Pupil dilation as an indicator of cognitive workload in human-computer interaction. In *Proceedings of the International Conference on HCI* (pp. 542-546).
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372–422.
- Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific Studies of Reading*, 10(3), 241–255.

- Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, *14*(3), 191-201.
- Rayner, K., & Pollatsek, A. (1997). Eye movements, the eye-hand span, and the perceptual span during sight-reading of music. *Current Directions in Psychological Science*, *6*(2), 49–53.
- Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). The effect of plausibility on eye movements in reading. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *30*(6), 1290–301.
- Rosemann, S., Altenmüller, E., & Fahle, M. (2016). The art of sight-reading: Influence of practice, playing tempo, complexity and cognitive skills on the eye-hand span in pianists. *Psychology of Music*, *44*(4), 1–16.
- Schluroff, M. (1982). Pupil responses to grammatical complexity of sentences. *Brain and Language*, *17*(1), 133–145.
- Sevilla, Y., Maldonado, M., & Shalóm, D. E. (2014). Pupillary dynamics reveal computational cost in sentence planning. *Quarterly Journal of Experimental Psychology*, *67*(6), 1041–52.
- Stewart, A. J., Pickering, M. J., & Sturt, P. (2004). Using eye movements during reading as an implicit measure of the acceptability of brand extensions. *Applied Cognitive Psychology*, *18*(6), 697–709.
- Truitt, F., Clifton, C., Pollatsek, A., & Rayner, K. (1997). The perceptual span and the eye-hand span in sight reading music. *Visual Cognition*, *4*(2), 143–161.

- Warren, T., & McConnell, K. (2007). Investigating effects of selectional restriction violations and plausibility violation severity on eye-movements in reading. *Psychonomic Bulletin & Review*, *14*(4), 770–775.
- Waters, A. J., Underwood, G., & Findlay, J. M. (1997). Studying expertise in music reading: Use of a pattern-matching paradigm. *Perception & Psychophysics*, *59*(4), 477–488.
- Wurtz, P., Mueri, R. M., & Wiesendanger, M. (2009). Sight-reading of violinists: Eye movements anticipate the musical flow. *Experimental brain research*, *194*(3), 445-450.

Table 1

Means (and Standard Errors) for Experiment 1 (without metronome)

Region	Pre-target	Target	Post-target
First-Pass Time (ms)			
Congruent	1607 (85)	1655 (97)	1798 (88)
Anomalous	1627 (75)	1579 (78)	1689 (110)
Total Time (ms)			
Congruent	1911 (88)	1932 (104)	2220 (119)
Anomalous	1943 (89)	1944 (94)	2186 (135)
Regressions Out (count)			
Congruent	.11 (0.02)	.11 (0.02)*	.35 (0.02)
Anomalous	.14 (0.03)	.17 (0.02)*	.36 (0.02)
Average Pupil Size (pa)			
Congruent	317 (22)	312 (22)	300 (21)*
Anomalous	316 (24)	313 (23)	304 (21)*
<i>Performance Analysis</i>			
Timing Errors (ms)			
Congruent	127 (14)	145 (14)	128 (14)
Anomalous	129 (14)	142 (13)	125 (16)
Pitch Errors (semitones)			
Congruent	0.10 (0.10)	0.07 (0.10)	0.03 (0.06)
Anomalous	0.12 (0.13)	0.06 (0.09)	0.01 (0.03)

* $p \leq .05$ ** $p \leq .01$

Table 2

Means (and Standard Errors) for Experiment 2 (with metronome)

Region	Pre-target	Target	Post-target
First-Pass Time (ms)			
Congruent	1765 (73)	1608 (49)	1837 (89)*
Anomalous	1662 (43)	1663 (54)	1676 (79)*
Total Time (ms)			
Congruent	1996 (66)	1845 (46)*	2201 (67)
Anomalous	1960 (34)	1981 (53)*	2141 (71)
Regressions Out (count)			
Congruent	.12 (0.02)	.08 (0.02)**	.34 (0.03)
Anomalous	.16 (0.02)	.19 (0.02)**	.37 (0.03)
Average Pupil Size (pa)			
Congruent	351 (14)	344 (13)*	334 (13)**
Anomalous	353 (14)	351 (13)*	345 (13)**
<i>Performance Analysis</i>			
Timing Errors (ms)			
Congruent	54.7 (7.6)	55.1 (8.3)	54.0 (8.1)
Anomalous	53.4 (7.7)	56.5 (8.1)	57.7 (9.3)
Pitch Errors (semitones)			
Congruent	0.13 (0.12)	0.10 (0.11)	0.06 (0.08)
Anomalous	0.15 (0.13)	0.13 (0.12)	0.05 (0.07)

* $p \leq .05$ ** $p \leq .01$

Figure 1. Example items in each time signature used in the experiment. Target bars are depicted by the overhead bracket. The two melodic contexts ([1] & [2] vs. [3] & [4]), and two versions of the target bar ([1] & [3] vs. [2] & [4]), are each presented in one congruent and one anomalous version.

Figure 2. Regions of interest in a sample stimulus. Each region was one musical bar.