

# Handling of incongruences in music notation during singing or playing

International Journal of  
Music Education  
1–21

© The Author(s) 2020

Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/0255761420944036  
journals.sagepub.com/home/ijm**Natalia Chitalkina**   
University of Turku, Finland**Marjaana Puurtinen**   
University of Turku, Finland**Hans Gruber**  
University of Regensburg, Germany  
University of Turku, Finland**Roman Bednarik**  
University of Eastern Finland, Finland

## Abstract

During music reading, performers create expectations of the upcoming music. When these expectations are violated due to changes in the notation, performers have to adjust their reading and adapt their motor responses to match this new information. In this study, we examine how selected background, outcome, and process measures reflect the successful handling of incongruences during music reading. Twenty-four performers were tasked with singing or playing versions of “Mary had a little lamb” in two different tonalities. Some versions contained a surprising element: a bar shifted down a tone. Selected outcome and process measures, such as performance accuracy and eye-movements during music reading (eye-time span, duration of first-pass fixations and pupil dilation), were analyzed. In sum, incongruence in music notation not only increased the number of performance mistakes, but incongruent melodies also led to micro-level changes in the reading processes. We propose that understanding the cognitive strategies for successful music reading requires going beyond the more traditional outcome measures and focus on the detailed analyses of the reading process itself.

## Keywords

Cognitive tests, eye tracking, incongruence, music reading, pupil size

---

## Corresponding author:

Natalia Chitalkina, University of Turku, 20500 Turku, Finland.  
Email: natalia.chitalkina@utu.fi

## **Introduction**

Western music notation is a symbolic language often applied in learning and performing on various musical instruments and singing. It includes a variety of symbols and complex musical concepts of higher abstraction levels (Chitalkina et al., 2019). It therefore comes as no surprise that many beginners experience difficulties while learning to read music. In fact, musically untrained beginners' assumptions about Western music notation are not even consistent with its real conventions (Tan et al., 2009). In contrast, experienced musicians can not only interpret the symbol system in detail, but it seems like they are driven by this knowledge, especially when they have to deal with its local complexities (Wolf, 1976).

A special case of such local complexities are unexpected note changes that do not fit the general structure of a music piece. Knowledge about how musically experienced individuals read these incongruences in music scores is useful for music education, because it might help to understand how they try to adapt their reading strategies. This knowledge includes different elements, such as performance accuracy, knowledge of performers' cognitive characteristics that might be associated with music reading, and knowledge of how the difficulties are solved during the actual reading process. A suitable experimental approach is required to take all these aspects into account. In the following sections we will briefly introduce a history of research about dealing with incongruence in music reading, with the purpose of informing the educational field about the development of the field as well as introduce our own, integrated approach.

## **Early work: the performance approach**

Research on how musicians perform incongruent music notation began already in the 1970s—and by accident. A musician Goldovsky (Wolf, 1976) discovered a misprint (a wrong single note in a chord symbol) in a music score by listening to how one of his students performed the piece. Subsequently, Goldovsky asked a group of experienced readers to find the misprint, allowing them to perform the piece as many times as needed. Surprisingly, none of the experts was able to find the mistake—they played the erroneous chord according to their expectations and not according to the (mis)printed edition.

Sloboda (1976) further extended this study by investigating the handling of incongruent notation in a more systematic way. He introduced deliberate misprints in four fragments of unknown pieces, equally spreading them among the staves and positions of a musical phrase. However, none of his experienced pianists, who played each piece twice, was able to find all changes.

In these early studies, the research focus was mainly on performance accuracy, that is, whether participants played according to misprinted music or whether they played according to their expectations of how the music should sound. Despite the relatively simple experimental designs, this approach led to the discovery of the proof-readers' error (Sloboda, 1976; Wolf, 1976): a phenomenon describing when experienced music readers ignore incongruent notes and play what they expect the notes to contain instead. The discovery confirms that experienced performers do have predictions about music scores and they actively apply them during music reading.

## **Focus on performers' cognitive characteristics**

The next generation of researchers followed a new research path and started to investigate the possible connections between music reading and performers' cognitive characteristics. Music reading and cognitive parameters were measured with special tests. Afterward, regression analysis was usually applied to search for significant predictors of music-reading skill. This line of research

mainly investigates what might affect sight-reading and sight-singing, that is, the task performing unfamiliar music scores “at first sight.”

Several types of cognitive processes have been suggested to be linked to sight-reading skill, among them text reading comprehension (Gromko, 2004), rhythmic audiation (Gromko, 2004), spatial orientation skills (Gromko, 2004; Hayward & Gromko, 2009), performing a repertoire of rehearsed music (McPherson, 1994), technical proficiency skills (Hayward & Gromko, 2009), speed of information processing (Kopiez & Lee, 2006, 2008), inner hearing (Kopiez & Lee, 2006, 2008), and psychomotor speed (Kopiez & Lee, 2006). Interestingly, reading comprehension was found (Reifinger, 2018) to be a significant factor in explaining variance in the beginner-level sight-singing, whereas reading fluency was not. These observations make the possible correlations between text and music reading worth of further investigation.

The role of working memory has also been under study (Herrero & Carriedo, 2019; Kopiez & Lee, 2006; Meinz & Hambrick, 2010). However, it seems that its significance varies depending on the difficulty of the music-reading task, increasing for easy tasks and decreasing for more complicated ones (Kopiez & Lee, 2006). Meinz and Hambrick (2010) claimed an incremental positive effect of working memory on sight-reading, and that this is independent from deliberate practice. Herrero and Carriedo (2019) defined the roles of different working memory sub-processes in music sight-reading: retrieval and transformation sub-processes contributed to music sight-reading independent from the level of task difficulty, but the substitution sub-process only contributed to music sight-reading in easy tasks.

The performers' cognitive characteristics approach extended the performance approach by investigating performer-related characteristics that might precede a successful music-reading performance. Thus, performers are not only able to anticipate notation, but they are also different in terms of their cognitive abilities and this might predict how they are able to read the music.

## **Introducing process measures to music-reading studies**

Despite advances in methodology and interesting findings, the two research approaches described above were not suitable for providing a deeper insight into the process of reading of incongruent music notation: scholars were only able to analyze the outcome parameters, such as correctness of the performances, but were not able to tap into the course of the reading process itself. Nowadays rapid advances in the development of the eye-tracking method enable the detailed study of how musicians cope with complexities in music notation (Madell & Hébert, 2008; Puurtinen, 2018). For instance, we can investigate how musicians prepare to perform incongruences and how they perform the subsequent bars (Hadley et al., 2018; Penttinen et al., 2015).

Eye-movement behavior is usually described in terms of fixations and saccades. Fixations might be defined as short “stops,” or suppressing of ocular drifts, when a steady retinal image of an object of interest is maintained (Komogortsev et al., 2010). Saccades are rapid movements from one fixation to another. The typical fixation duration for music reading has been reported to be around 500–700 ms (Arthur et al., 2016; Penttinen et al., 2015). However, it might be influenced by musician-related factors, such as music-reading expertise (Puurtinen, 2018). Fixations might be further divided into first-pass and second-pass fixations (Holmqvist et al., 2011). First-pass fixations are fixations made during the first entrance of the specific area until the first exit of this area. Second-pass fixations are fixations during a revisit to the area in question.

Scholars interested in how experienced participants read incongruent music scores have differed on whether they apply familiar or unfamiliar music pieces or probes, into which the incongruences were inserted. The first line of research investigated how musicians process incongruent changes in unfamiliar notation (Ahken et al., 2012; Hadley et al., 2018). Ahken et al. (2012)

explored how pianists sight-read musical phrases containing incongruences that violated the musical tonality. The study (Ahken et al., 2012) provided evidence that these kind of incongruences might cause changes in eye-movement behavior, such as an increase in mean proportion and duration of fixations. Hadley et al. (2018) studied how pianists sight-read unfamiliar melodies that contained an anomalous, or “not typical,” bar. In this study, the pupil size parameter associated with cognitive arousal and effort (Holmqvist et al., 2011) was apparently analyzed for the first time in music-reading studies. Although experienced pianists were able to integrate notes they read into the prior context, anomalous pitch relationships led to processing difficulties that were reflected in eye-movement parameters through an increase in total duration of fixations and a decrease of first-pass fixation duration in the bar after the incongruence (Hadley et al., 2018). The first-pass fixation duration might have reflected compensatory processes while dealing with incongruence, that is, decreased duration of first-pass fixations immediately after the incongruent bar (Hadley et al., 2018). In addition, Hadley et al. (2018) also demonstrated that the mean pupil size could be a useful parameter for studying processing difficulties in music reading, as the mean pupil size might increase immediately after the incongruent section.

The second line of eye-tracking research about the handling of local complexities in music notation (Penttinen et al., 2015) considers that a different research design is needed: incongruent changes should be inserted into familiar music scores. Research from the music perception (Gunter et al., 2003; Vuust et al., 2018) field provides a solid ground for this approach by suggesting that incongruence detection in music relies on modality-independent processing mechanisms, and unpredictability of music features makes deviation from prediction less salient.

Penttinen and colleagues (2015) applied eye-tracking in their study comparing skilled amateurs’ and professional-level musicians’ performances of a familiar melody, with and without incongruent bars, and during a temporally controlled music-reading task. This study focused on errorless performances, that is, proof-readers’ errors were excluded from the analysis. Although the two groups of participants generally differed according to their eye-movement parameters, the study suggested that these differences disappeared when musicians performed the incongruent parts of the notation. The study also applied the eye-hand span measure for the first time in the study of incongruences in music notation. This parameter that reflects how far ahead the gaze is from the current point of execution was created to analyze time-limited visual-motor coordination in music reading (Sloboda, 1974). In the study by Penttinen and colleagues (2015), it was defined as the time between the performed note and the concurrent point of gaze. Penttinen et al. (2015) demonstrated that more experienced musicians tend to operate with shorter average fixation durations and apply longer eye-hand spans more often, but these differences vanished at the face of incongruent parts of notation.

This study replicates and expands from this prior study by exploring if incongruence effects appear when the performer performs the music in two different tonalities and two different modalities. We will focus on three eye-movement parameters: two that have already proven useful in previous studies, first-pass fixation duration and mean pupil size in first-pass fixations, and one new parameter, the eye-time span (ETS; Huovinen, Ylitalo and Puurtinen, 2018).

The ETS parameter represents an improved version of the above-mentioned eye-hand span. Huovinen et al. (2018) pointed out two important limitations of the eye-hand span parameter. First, the use of the “hand,” that is, currently executed note as a starting point of the measurement, is not very suitable for the study of music—structural complexities. Second, the parameter might be imprecise, because the fixation typically occurs slightly earlier or later than the execution of the note. Instead, the ETS is a distance between the time a fixation is initiated and the ongoing musical time at the onset of this fixation. Huovinen and colleagues (2018) indicated that the appearance of even slight complexities in music notation (such as larger melodic intervals) might increase the ETS. This parameter has not been previously applied in studies about the handling of incongruences in music notation.

All in all, the eye-tracking studies described above provided insight into the actual process of music reading. They also established eye-movement parameters relevant for handling incongruent music and opened a discussion around music material that should be used as stimuli in its study.

## Unaccounted aspects of incongruence in music reading

Despite the many different aspects of handling local incongruences in music scores that have been studied, the stability of individuals performers' strategies across related but slightly different performance tasks is yet understudied. In our study, we address this issue by inviting our participants to perform in two different modalities (singing, playing) as well as in two different tonalities.

To begin, how well musicians are able to cope with incongruences might depend on the way musicians perform the music. For instance, when a musician plays on a musical instrument, they should not only understand the score, but also convert or compile musical pattern to the "language" of the musical instrument (i.e., keys or strings) as well as plan for the correct motor sequences (Chitalkina et al., 2019). In singing, musicians also have to plan the motor responses, but they are executed by different kinds of physiological activity. The compilation phase during singing is reduced, since singing does not include an interaction with any external instrument. However, notational audiation has been shown to be more important for sight-singing than for sight-reading on a piano (Fine et al., 2006), which can make sight-singing more complicated than sight-playing.

Another aspect is how musicians deal with incongruences in the same melody transposed into different tonalities. The concept of tonality is one of the most important concepts of music theory. Many students have difficulties while sight-reading music in different tonalities (see Alexander & Henry, 2012). Abstract knowledge of the tonality concept is acquired with sufficient skill level. When musicians have to perform from incongruent notation in different tonalities, they not only have to rapidly integrate incongruent parts in the previous musical context, but also apply their concept of tonality. The only eye-tracking study addressing tonality (Ahken et al., 2012) was conducted with some limitations in the study design: for instance, the melodies in congruent and incongruent conditions of the study were different, hindering direct comparisons. Furthermore, incongruent changes were inserted in the last bar of the melodies making it impossible to study the cognitive processing immediately after the incongruent part of the notation. In sum, an interesting issue that has not been investigated yet is how fluently musicians apply their strategies of coping with incongruences in the same music piece given in different tonalities, or whether the fact that musicians produce the sounds by themselves (singing) or through an instrument influences the fluency of their coping strategies.

## Aim and hypotheses

The main aim of this study was to explore how experienced music readers deal with local incongruences in familiar music in two different tonalities and performed in two different modalities (playing or singing). Instead of selecting one approach, we will use an integrated approach, combining the suggestions of the research traits described above, and conduct a careful investigation of both outcome and process parameters. In this study, we have chosen to insert incongruent changes in a familiar melody.

We build our study on the following hypotheses. (1) We hypothesize that there is a significant correlation between the number of correct performances of incongruent melodies and selected cognitive characteristics, namely, text reading ability and short-term memory. (2) We hypothesize that incongruent changes in familiar music notation will lead to processing difficulties that are reflected (a) in the common reading strategy suitable for the congruent music notation and (b) in

the increased number of incorrect performances of incongruent melodies. (3) We hypothesize that incongruences increase processing difficulties more in the more complicated tonality. (4) We hypothesize that the incongruences increase processing difficulty more when musicians are playing the piano, compared to when they are singing.

## **Method**

### *Participants*

A total of 24 musically experienced participants with normal or corrected to normal vision participated in the study. The age of the participants varied between 16 and 57 years ( $M=32.50$ ,  $SD=10.18$ , 20 females). Nine participants had 20 or more years of music experience, eight participants had more than 10 years of music experience and seven participants had 4–7 years of music experience. Two participants were excluded from the data analyses: one participant was not able to perform in time with a metronome, and the another participant was excluded because of problems with calibration.

### *Tests and measures*

In addition to the eye-tracking data, we also measured the number of correct performances, the speed and accuracy of text reading, and auditory short-term memory capacity of the study participants. The performance data was audio- and video recorded for the subsequent analysis, conducted by two researchers.

Reading abilities of participants were measured with the help of words' and pseudowords' reading tests from the Assessment Battery for Reading Disabilities in Young and Adults (Nevala et al., 2007). The reading tests were administered in Finnish language. Participants were asked to read the list of words and pseudowords. Speed and accuracy of reading were assessed.

A digit span memory test (taken from the Dyslexia and Literacy International Charity website) was administered to assess auditory short-term memory of the participants. Participants were asked to repeat increasing rows of numbers read on tape by a native Finnish speaker either in direct or in reverse order.

### *Stimuli*

The music notation of the folk song "Mary had a little lamb" (composed by L. Mason/Public Domain) was presented on a white computer screen with the resolution of  $1,920 \times 1,080$  pixels. Participants were seated approximately 65 cm away from the screen. The original melody was presented in C major and transposed by the authors to a more complicated tonality containing five sharps, that is, B major.

In this study, we used the two variations of the original song from the study by Penttinen et al. (2015). Both of these variations included one bar shifted down, that is, an incongruent bar (see Figure 1). These incongruent variations were further transposed to B major (see Figure 2). Music scores of all melodies were prepared in the Sibelius music notation software.

### *Apparatus*

All music performances were video recorded with two cameras: one camera was adjusted to capture the computer screen, and the other one was put in the corner of the laboratory. Participants

Original

pre a pre b tar a tar b post a post b

Variation 1

pre a pre b tar a tar b post a post b

Original

pre a pre b tar a tar b post a post b

Variation 2

pre a pre b tar a tar b post a post b

**Figure 1.** “Mary had a little lamb” in C major and its two variations with areas of interests (AOIs) applied in the data analysis.

Original

pre a pre b tar a tar b post a post b

Variation 1

pre a pre b tar a tar b post a post b

Original

pre a pre b tar a tar b post a post b

Variation 2

pre a pre b tar a tar b post a post b

**Figure 2.** “Mary had a little lamb” in B major and its two variations with areas of interests (AOIs) applied in the data analysis.

performed on a Yamaha electric piano. The eye-movement data were recorded with a remote binocular Tobii TX 300 Eye-Tracker with a recording frequency of 300 Hz. Fixations were classified according to the Velocity-Threshold Identification fixation classification algorithm (Olsen, 2012) of the Tobii Pro Studio. Luminance in the room was kept the same during the whole study. In

**Table 1.** Study design (o—original, v—variation).

Set						
1	Singing	Co Cv1	Bo Bv2	Playing	Bo Bv1	Co Cv2
2	Singing	Bo Bv1	Co Cv2	Playing	Co Cv1	Bo Bv2
3	Playing	Co Cv2	Bo Bv1	Singing	Bo Bv2	Co Cv1
4	Playing	Bo Bv2	Co Cv1	Singing	Co Cv2	Bo Bv1

addition, the colors of calibration dots and background were identical to the colors of the stimuli and background in the study for accurate pupil size measurements.

### Procedure

In the beginning of the study participants were informed about the aims and they signed an informed consent form. Participants were introduced to the experimental setting and adjusted the chair to a convenient height to be able to perform on an electric piano with their right hand. After that, the eye-tracking study with three parts was conducted. First, the participants could play and then sing “Mary had a little lamb” in both C and in B major in time with a metronome set at 60 bpm, while reading the music from the computer screen. The researcher synchronized the change of the slides with the metronome. Participants were asked to play on a piano the melodies using their right hand only. In addition, participants were told to sing the melodies using the syllable “la” instead of words of the song. All melodies were sung without any instrumental accompaniment.

Next, participants were informed that some of the melodies would contain some alterations from the original melody and that they should perform according to music scores they would see on the screen. Each participant performed one out of four sets of stimuli in time with a metronome set at 60 bpm (see Table 1). Two calibrations were conducted, one prior to singing and another prior to playing. Finally, participants filled in a background questionnaire including questions about their previous music experience and self-evaluation of sight-reading and sight-singing skills. After the words’ and pseudowords’ reading test and the digit span memory test, an informal discussion of the participant’s reading process and looking at the eye-movement recordings ended the measurement session. The whole experiment lasted around 1 hour.

### Data analysis

**Performance data.** The first author evaluated the correctness of playing and singing performances with the help of video and audio recordings, and borderline cases were evaluated jointly by the first and the second authors (the latter being a professional musician). The analysis focused on the correctness of the performance of the altered bar, the bar preceding the altered one and the bar after the altered in variations and corresponding bars of the originals. Performances with errors were included in the analysis. The number of mistakes in four variations of the melody (playing in C major and in B major, singing in C major and in B major) was calculated for each participant.

**Reading skill and working memory.** Time (in seconds) spent on reading words, pseudowords, and number of mistakes in pseudowords were calculated to analyze the speed and accuracy of participants’ reading. In order to analyze the capacity of the auditory short-term memory, the



number of correctly recalled numerals in direct order and the number of correctly recalled numerals in reverse order were calculated as well as their sum (the total number) and the standard score for the total number. Nonparametric Kendall's tau  $b$  tests were conducted in SPSS to analyze the correlation between the number of mistakes in variation performances and each of the following cognitive parameters: (1) time spent on reading words, (2) time spent on reading pseudowords, (3) number of mistakes in pseudowords, (4) the number of correctly recalled numerals in direct order, (5) the number of correctly recalled numerals in reverse order, (6) the total number of recalled numerals, (7) the standard score for the digit span test.

*First-pass fixation time, ETS, and pupil size.* The aim of the analysis was to analyze both perfect performances and performances containing mistakes in and around the space that (in four of the melodies) included an altered bar. In order to be able to analyze processing of incongruence in more details and to see how and where changes in the coping strategy occur, we decided to choose a half of the bar as a unit of our analysis. We created six squared areas of interest (AOI) around half-bars of the altered bar, one preceding bar, and one following bar in the variation melodies as well as around the corresponding half-bars of the original melodies using the Tobii Pro Studio software: pre-target a, pre-target b, target a, target b, post-target a, post-target b (see Figures 1 and 2). If there were no fixations on the AOI, the data was handled as missing in the analysis. The first-pass fixation duration and ETS were calculated for each AOI half-bar with the custom written scripts in R. The ETS was calculated according to the procedure described by Huovinen and colleagues (2018). First, the ETS for fixations that occurred exactly on the first note of the beat was measured by calculating the difference between the time (A) of the AOI beat (musical time when the first note of the beat should be executed) and the time (B) of the first fixation on AOI (time when the participant looked at the note for the first time):  $(ETS=A-B)$ . Some fixations occurred to the left or to the right of the first note of the beat, and if so, we also calculated how many pixels the fixation was to the left or to the right from the beat note (C) as well as how long one pixel change took (D). For these fixations, we measured ETS in the following way:  $(ETS=A+C*D-B)$ . The mean pupil size during first-pass reading of an AOI was calculated with the custom written scripts in R for four of the above-mentioned AOIs only: second half of the pre-target, both target half-bars, and the first post-target half-bar. The first-pass reading includes all fixations targeting the AOI before leaving it for the first time (typically —one to two fixations).

*Statistical analyses.* In addition to the analysis of the performance quality, we performed Pearson's chi-square tests to investigate the relationship between each of the following parameters: (1) type of the melody, (2) modality of performance, (3) tonality of performance and correctness of performance. Furthermore, to analyze the relationship between cognitive characteristics and reading processes of incongruent melodies, we conducted nonparametric Kendall's tau correlation tests.

The relationship between modality, tonality, and congruence in each of the parameters—(1) mean pupil size in first-pass fixations, (2) duration of first-pass fixations, (3) ETS—was analyzed with linear mixed models in the lme4 package (Bates et al., 2014) of R. We entered the following fixed effects into the model: mistake (if a mistake was made), congruence, tonality, modality as well as two-way interactions modality\*tonality, modality\*congruence, and tonality\*modality and a three-way interaction tonality\*modality\*congruence. As for the random effects, we entered intercepts for subjects, as well as by-subject random slopes for the effect of mistakes. In the analysis, we always focused on the interpretation of the highest-order interaction in the case it was significant.

**Table 2.** Performance quality.

	Singing		Playing	
	Correct	Incorrect	Correct	Incorrect
C congruent	22	0	21	1
C incongruent	13	9	17	5
B congruent	22	0	18	4
B incongruent	11	11	15	7
Total	68	20	71	17

## Results

### *Incongruent music performance and cognitive characteristics of performers*

In this section, we present results concerning the relationship between performer-related cognitive characteristics and the number of correct variation performances and correctness of music performances.

*Performance quality and task type.* First, performing of incongruent melodies caused difficulties for the participants: Pearson's chi-square analysis showed a significant association between the type of the melody (congruent vs. incongruent) and the correctness of the performance,  $\chi^2(1)=24.95$ ,  $p < .001$ , odds ratio (OR)=9.49. Despite the familiarity of the song, the study task was not overly simple for the participants and therefore allowed to explore the handling of incongruent parts of the music notation.

As can be seen from Table 2, participants seemed to sing original melodies slightly better than they played them; however, Pearson's chi-square analysis did not reveal any significant association between the modality of the performance (singing vs. playing) and the correctness of the performance,  $\chi^2(1)=3.14$ ,  $p=.076$ . In addition, there was no significant correlation between the tonality of the performance (C major vs. B major) and the correctness of the performance,  $\chi^2(1)=0.79$ ,  $p=.375$ .

*Performance quality, reading skill, and working memory.* In order to analyze whether there was any significant relationship between our selected cognitive characteristics (performer-related factors) and correctness of reading of incongruent melodies, we carried out nonparametric Kendall's tau correlation tests.

Unlike in some prior reports, the analysis did not indicate any significant relationship between the reading or short-term memory parameters and the number of mistakes in variation performances (see Table 3). Table 4 reports on cognitive tests on memory and reading among the experimental group.

### *Process-measures parameters*

To get an insight into the music-reading processes, when coping with incongruences, we analyzed three parameters: ETS, first-pass fixation duration, and mean pupil size for first-pass fixations.<sup>1</sup>

*ETS.* At the beginning, we analyzed how participants approached congruent and incongruent half-bars in the melodies. Assuming a normal distribution, we used linear mixed models to analyze the relationship between the ETS and tonality, modality, and congruence. The fitted linear mixed model for the pre-target a half-bar (see Table 5) showed significant effect of congruence,<sup>2</sup>

**Table 3.** Correlations between the number of mistakes and cognitive parameters.

Cognitive test	$\tau$	$p$
Memory parameters		
Digit forwards raw	-0.02	0.91
Digit backwards raw	-0.01	0.95
Digit total raw score	-0.03	0.88
Digit total standardized score	-0.05	0.79
Reading parameters		
Time spent on reading words	0.20	0.23
Time spent on reading pseudowords	0.09	0.58
Number of mistakes in pseudowords	-0.26	0.14

**Table 4.** Results of the standardized test scores for the study participants.

Cognitive test score	$M$	$SD$
Memory parameters		
Digit forwards raw	9.77	2.09
Digit backwards raw	8.23	2.45
Digit total raw score	18	4.32
Digit total standardized score	100.32	15.79
Reading parameters		
Time spent on reading words	19.36	4.85
Time spent on reading pseudowords	37.95	10.51
Number of mistakes in pseudowords	2.55	1.74

$SD$ : standard deviation.

$\chi^2(1)=14.41, p<.001$ . Congruence (performing of variations) increased the duration of ETS by about  $334.47 \pm 85.15$  ms. Thus, participants approached the first part of the pre-target bar with longer spans in incongruent melodies. This finding is in line with the previous research (Huovinen et al., 2018) suggesting that the ETS effect of upcoming difficulties might be registered on the music notation symbols preceding the most complex symbols.

The fitted linear mixed model for the target b half-bar (see Table 5) showed a significant two-way modality\*tonality interaction,  $\chi^2(1)=5.82, p=.02$ , and a significant effect of congruence,  $\chi^2(1)=4.09, p=.04$ . Participants (Figure 3) tend to operate with lower ETS when they had to sing in a more complicated key of B major than when they had to play in B major. It is possible that participants needed more time to translate the notation into “the piano language” and plan the motor sequences in B major. Congruence (performing of variations) lowered ETS by  $158.59 \pm 77.56$  ms. The effect of congruence might reflect processing difficulties: participants might not have enough resources to approach the second half of the incongruent bar as early as they reach the corresponding half-bar of the original melody, because they have already started to cope with the incongruent notation. Interestingly, there were no significant interactions of congruence with other parameters.

The fitted linear mixed model for the post-target a half-bar showed a significant two-way modality\*tonality (Figure 4) interaction,  $\chi^2(1)=9.31, p=.002$ . Similar to the target b half-bar, participants had lower ETS when they sang in B major than when they played in B major.

**Table 5.** Parameter estimates of the final models fitted for ETS.

	Estimate	SE	T
Pre-target a			
(Intercept)	1,460.06	132.09	11.05
Congruence (incongruent)	334.47	85.15	3.93
Target b			
(Intercept)	1,279.72	111.28	11.50
Modality (singing)	-325.70	111.25	-2.93
Tonality (C major)	-117.30	108.95	-1.08
Congruence (incongruent)	-158.59	77.56	-2.05
Modality (singing):Tonality (C major)	375.67	153.29	2.45
Post-target a			
(Intercept)	1,611.51	164.86	9.78
Modality (singing)	-318.40	123.17	-2.59
Tonality (C major)	-53.84	123.51	-0.44
Congruence (incongruent)	-105.87	94.34	-1.12
Mistake (correct)	-170.88	183.98	-0.93
Modality (singing):Tonality (C major)	526.79	169.56	3.11
Post-target b			
(Intercept)	1,303.09	114.13	11.42
Modality (singing)	-415.34	107.76	-3.85
Tonality (C major)	-419.95	138.50	-3.03
Tonality (B major):Congruence (incongruent)	-78.42	111.16	-0.71
Tonality (C major):Congruence (congruent)	390.55	109.69	3.56
Mistake (correct)	94.76	118.33	0.80
Modality (singing):Tonality (C major)	416.71	148.83	2.80

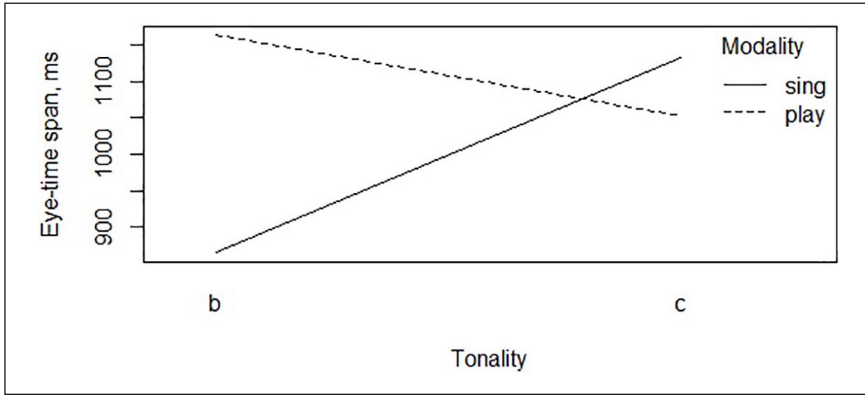
SE: standard error.

The fitted linear mixed model for the post-target b half-bar (see Table 5) showed a significant two-way tonality\*congruence (Figure 5) interaction,  $\chi^2(1)=4.41$ ,  $p=.04$ , and a significant two-way modality\*tonality (Figure 6) interaction,  $\chi^2(1)=7.58$ ,  $p=.006$ .

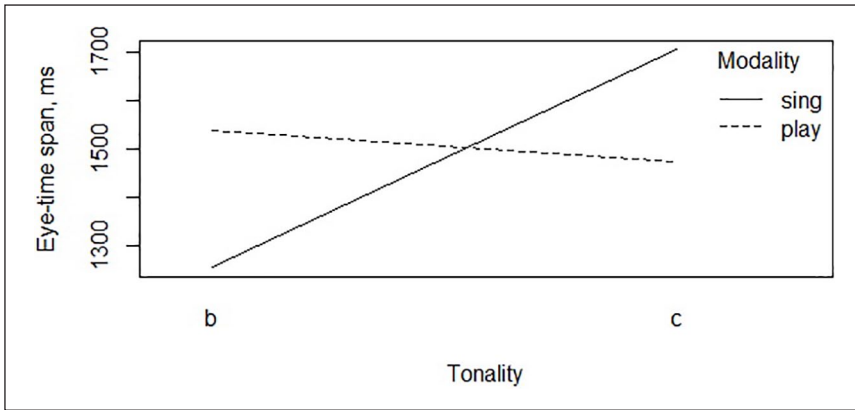
Participants had longer ETS when they performed the second half of the post-target bar of incongruent melodies in B major compared to C major. Interestingly, participants tend to approach this half-bar both in original and incongruent melodies with ETS of the similar length in B major. It seems that a difficult tonality might require more planning. In contrast, participants operated with shorter ETS performing this half-bar in incongruent melodies in C major compared to congruent melodies in C major. Similar to the target b and post-target a half-bars, participants operated with lower ETS when they sang in B major than when they played in B major likely due to the increased cognitive demands of playing in a more complicated key.

*First-pass fixation duration.* Using linear mixed models, we analyzed the relationship between the duration of first-pass fixations and tonality, modality, and congruence. The fitted linear mixed model for the pre-target b half-bar (see Table 6) showed a significant two-way interaction (Figure 7) between tonality and congruency,  $\chi^2(1)=8.41$ ,  $p=.004$ .

Thus, participants made longer first-pass fixations to the half-bar immediately preceding the incongruent half-bars in B major compared to C major. It seems that it might be more difficult to prepare for the upcoming incongruence in B major.



**Figure 3.** Mean eye-time spans measured for singing and playing performances in C and B majors in the target b half-bar

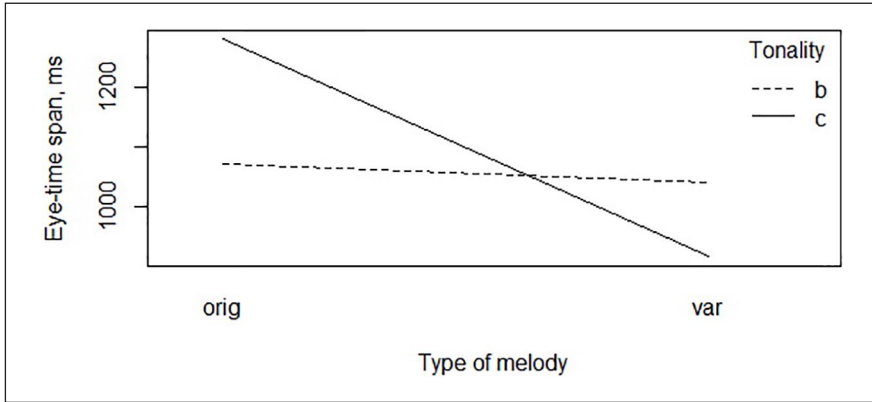


**Figure 4.** Mean eye-time spans measured for singing and playing performances in C and B majors in the post-target a half-bar.

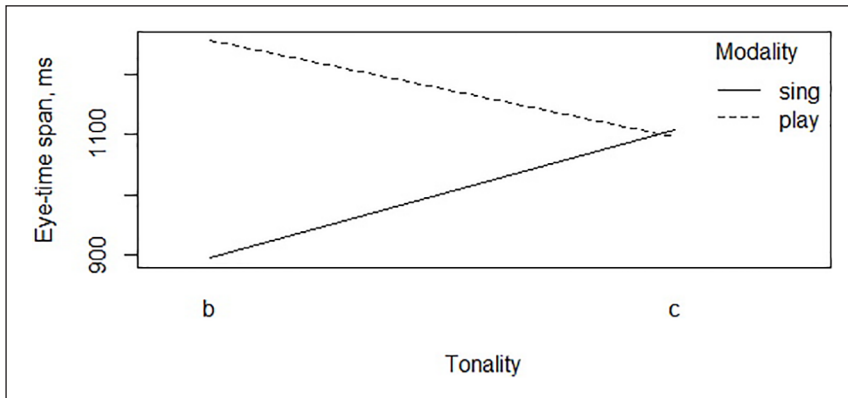
The fitted linear mixed model for the target a half-bar (see Table 6) showed a significant effect of congruence,  $\chi^2(1)=7.35, p=.007$ . Congruence (performing of variations) lowered the duration of first-pass fixations by about  $199.95 \pm 72.41$  ms. The fitted linear mixed model for the post-target a half-bar (see Table 6) showed a significant effect of congruence,  $\chi^2(1)=4.79, p=.03$ . Congruence (performing of variations) increased the duration of first-pass fixations by about  $144.89 \pm 64.37$  ms. Interestingly, no significant interactions of congruence with other parameters were found in target and post-target half-bars.

Therefore, participants approached the first incongruent half-bar with shorter first-pass fixations and the first half-bar after incongruent half-bars with longer first-pass fixations compared to original melodies.

*Pupil size at first-pass fixations.* Assuming a normal distribution, we used linear mixed models to analyze the relationship between the mean pupil size in first-pass fixations and tonality, modality, and congruence. The fitted linear mixed model for the pre-target b half-bar (see Table 7) showed



**Figure 5.** Mean eye-time spans measured for original and variation melodies in C and B majors in the post-target b half-bar.

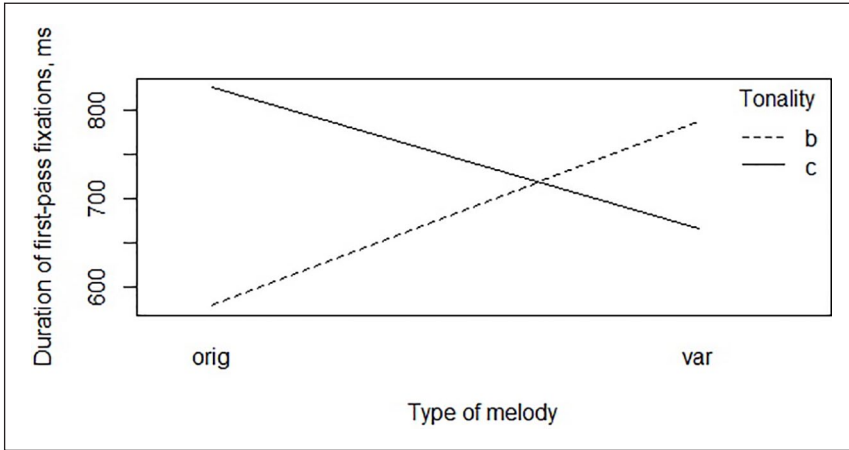


**Figure 6.** Mean eye-time spans measured for singing and playing performances in C and B majors in the post-target b half-bar.

significant effects of tonality,  $\chi^2(1)=6.88$ ,  $p=.009$ , and modality,  $\chi^2(1)=8.41$ ,  $p=.004$ . Tonality (performing in C major) lowered the mean pupil size during first-pass fixations by about  $0.05 \pm 0.02$  mm. Modality (singing) lowered the mean pupil size during first-pass fixations by about  $0.06 \pm 0.02$  mm.

The fitted linear mixed model for the target a half-bar (see Table 7) showed significant effects of tonality,  $\chi^2(1)=10.35$ ,  $p=.001$ , and modality,  $\chi^2(1)=9.95$ ,  $p=.002$ . Tonality (performing in C major) lowered the mean pupil size during first-pass fixations by about  $0.05 \pm 0.02$  mm. Modality (singing) lowered the mean pupil size during first-pass fixations by about  $0.05 \pm 0.02$  mm.

The fitted linear mixed model for the target b half-bar (see Table 7) showed significant effects of congruency,  $\chi^2(1)=37.3$ ,  $p<.001$ , modality,  $\chi^2(1)=7.75$ ,  $p=.005$ , and mistakes,  $\chi^2(1)=7.63$ ,  $p=.006$ . Performing with mistakes lowered the mean pupil size during first-pass fixations by about  $0.12 \pm 0.04$  mm. Modality (singing) lowered the mean pupil size during first-pass fixations by about  $0.06 \pm 0.02$  mm. Congruency (performing variations) increased the mean pupil size during first-pass fixations by about  $0.16 \pm 0.02$  mm.



**Figure 7.** Mean duration of first-pass fixations measured for original and variation melodies in C and B majors in the pre-target b half-bar.

**Table 6.** Parameter estimates of the final models fitted for the duration of first-pass fixations.

	Estimate	SE	t
Pre-target b			
(Intercept)	577.71	88.17	6.55
Modality (singing)	-0.14	62.94	0
Tonality (B major):Congruence (incongruent)	229.08	94.52	2.42
Tonality (C major):Congruence (congruent)	245.20	88.26	2.78
Tonality (C major):Congruence (incongruent)	104.01	93.91	1.11
Mistake (correct)	-110.29	99.09	-1.11
Target A			
(Intercept)	846.39	60.96	13.89
Congruence (incongruent)	-199.95	72.41	-2.76
Post-target a			
(Intercept)	736.46	59.33	12.41
Congruence (incongruent)	144.89	64.37	2.25

SE: standard error.

The fitted linear mixed model for the post-target a half-bar (see Table 7) showed significant effects of congruency,  $\chi^2(1)=33.28, p < .001$ , and modality,  $\chi^2(1)=7.24, p=.007$ . Modality (singing) lowered the mean pupil size during first-pass fixations by about  $0.06 \pm 0.02$  mm. Congruency (performing variations) increased the mean pupil size during first-pass fixations by about  $0.14 \pm 0.02$  mm.

Thus, the pupil size analysis indicated that singing from music scores might be a less challenging cognitive task, when compared to playing the piano from sheet music. It also seems that this was not affected by incongruence coping strategies, since the pupil size was lower for singing performances in all analyzed half-bars and there was no significant interaction between modality and congruence. Interestingly, the significant effect of tonality was found in the first two analyzed half-bars. It is generally known that the changes in pupil size parameter occur relatively slow, so

**Table 7.** Parameter estimates of the final models fitted for the mean pupil size in first-pass fixations.

	Estimate	SE	T
Pre-target b			
(Intercept)	2.83	0.06	45.25
Modality (singing)	-0.06	0.02	-2.95
Tonality (C major)	-0.05	0.02	-2.67
Target A			
(Intercept)	2.87	0.07	43.07
Modality (singing)	-0.05	0.02	-3.22
Tonality (C major)	-0.05	0.02	-3.30
Target b			
(Intercept)	2.86	0.07	40.72
Modality (singing)	-0.06	0.02	-2.84
Congruence (incongruent)	0.16	0.02	6.67
Mistake (correct)	-0.12	0.04	-3.21
Post-target a			
(Intercept)	2.86	0.07	41.13
Modality (singing)	-0.06	0.02	-2.83
Congruence (incongruent)	0.14	0.02	6.20

SE: standard error.

the pupil size measured in target a half-bar might actually reflect the pupil size before the target bar. It might be further confirmed by the occurrence of the effect of congruence only in the second half of the target bar. Therefore, it seems that performing in C major is less challenging, but only when the notation does not contain unexpected changes. The significant effect of tonality disappeared while participants coped with incongruence in the notation. Moreover, it seems that the increase in pupil size parameter might reflect difficulties in the cognitive processing of unexpected notes. However, we did not find any significant interaction of congruence with other parameters.

## Discussion

This study provides insight into how experienced music readers handle local incongruences in music, when they play and sing familiar music in two different tonalities. To investigate this issue, we applied an integrated approach, combining the approaches of prior researchers, and analyzed two types of measures: outcome and process measures.

To begin, we studied the relationships between correct performances of incongruent melodies and performers' cognitive characteristics. However, no significant relationship between these parameters appeared. Thus, it seems that at least with our group size of 24 musicians, text reading skill or short-term memory capacity was not associated with performance accuracy. This finding is in line with Reifinger (2018), who also has not found a connection between reading fluency and sight-singing. As expected, we found that incongruence changed the strategy in reading congruent melodies. We also noted an increase in the number of mistakes for incongruent melodies, compared to congruent ones. This finding is consonant with Fine et al. (2006), who provided evidence that incongruences increased errors in sight-singing.

However, by analyzing only performance accuracies, it was not possible to detect what was different in the reading of incongruent melodies compared to congruent, but process measures were



needed. Indeed, the ETS analysis revealed that reading of incongruent notation was different from reading of congruent notation, and the difference already appeared in the bar preceding the incongruent part, rather than exactly at that section. Thus, participants approached the first half of the pre-target bar with longer ETS in incongruent melodies, compared to congruent ones. The location of our ETS effect is in line with Huovinen and colleagues (2018), who reported that ETS increased at the symbols preceding their “more complex” notes. Therefore, it is possible that the preparation for handling the incongruence in the melody requires additional processing time, and this leads to the slight local lengthening of the ETS. In addition, musicians approached the second half of the target bar with lower ETS in incongruent melodies compared to congruent ones. This finding could signal processing difficulties for performing the incongruent section: starting to deal with the incongruence, they might not have enough time to approach the second half of the target bar as early on as they did in congruent melodies (without such local complexities).

In contrast to Hadley et al. (2018), we did observe effects in the first-pass fixation durations during the reading of the incongruent melodies. In our case, first-pass fixation duration decreased in the first half of the target bar and increased in the first part of the post-target bar. A possible explanation is that musicians might be already prepared to deal with the incongruence in the first half of the target bar, because their coping strategies were activated earlier (in the first half of the pre-target bar), as indicated by the increase of the ETS. Therefore, they might not need additional first-pass time for the first half of the target bar. However, in line with Hadley et al. (2018), an increase in pupil size was found in the second half of the target bar and the first half of the post-target bar. Since changes in pupil size occur with a delay, this effect might be attributable to difficulties in cognitive processing of the incongruent bar.

Tonality also affected the performances: participants performed C major variations better, compared to B major variations. Interestingly, on the eye-movement level, the tonality-related differences appeared before and after the incongruent bar. Indeed, participants fixated the second half of the pre-target bar with longer first-pass fixations for incongruent melodies in B major compared to incongruent melodies in C major. It therefore seems like the preparation for coping with incongruence in B major might be more difficult compared to C major. In addition, participants tended to approach the second half of the post-target bar in incongruent B major melodies with longer ETS compared to incongruent C major melodies.

The pupil size analysis did not reveal significant interactions between congruence and tonality. Instead, pupil size lowered for C major compared to B major performances around the beginning of the incongruent section. Due to the delay in pupil size change occurrence, these effects are attributable to the processing of the pre-target bar, and not the incongruent section itself. Thus, it seems that for our participants, at least, performing in B major was generally more cognitively demanding than performing in C major, which is consonant with Alexander and Henry’s (2012) finding of an increase in sight-reading errors for more complicated tonalities.

Comparing singing to playing, against our hypotheses, participants better played incongruent melodies than sang them. Moreover, there were no significant interaction between congruence and tonality with respect to our eye-movement measures. However, pupil sizes indicated that the processing difficulty of singing might be lower than when playing (from notation). Moreover, a significant interaction between modality and tonality was observed in the second half of the target bar and in the post-target bar, indicating that performers approached these half-bars with lower ETS during singing in B major compared to playing in B major. The opposite difference in ETS between playing and singing in C major was smaller. Thus, it seems like singing from notation might still be less cognitively demanding than playing, when process measures are taken into account. This finding might signal about the role of the motor component in the piano performances, when performers might need additional resources to plan for the motor sequences.

The pupil size analysis is presumably becoming a common tool in the music perception field. It therefore might be interesting to compare our results with the results from related domains. Recent research (Weiss et al., 2016) reported that when listening to vocal and piano music, pupil dilation was greater for vocal melodies than for piano melodies, probably due to enhanced arousal to vocal melodies. Comparing this to our opposite observation, the effects of the pupil size measure for singing and playing might be different for music reading and listening. Interestingly, another study (Liao et al., 2018) from the music perception field suggested that the mean pupil diameter is larger at surprising than at unsurprising moments in music. Thus, the pupillary dilation might reflect surprise in both music reading and music perception. This might provide another explanation why performing of incongruent melodies was more difficult in B major: It is possible that performing in B major is not common for this song and contained an element of “surprise,” of something uncommon. It differs from the tonality in which participants remember the song. In this sense, the increase in mean pupil size might signal about the use of the unusual tonality.

This study has some limitations in its design. First, the sample size of 24 subjects was rather small. Although all participants were experienced in music, the sample included professional musicians as well as non-professional performers. Therefore, further research conducted with a larger number of participants with equal professional level is needed to confirm our findings and add to them the potential effects of musical expertise or music-reading skill (Puurinen, 2018). In addition, during piano performances participants might look down at the piano keyboard that had a different luminance level compared to the luminance level of the computer screen with music notation that might affect the accuracy of the pupil size measurement. Thus, future studies should develop research designs taking also this issue into account. Furthermore, our analysis of performance accuracy focused only on the correctness of the incongruent parts of the variations and the corresponding parts of the originals, which might have affected the results.

Nevertheless, the integrated approach taken in this study combines a controlled design with an adequate level of ecological validity, which we believe to be good starting point for designing music-reading experiments. A further move toward more controlled designs might create highly unnatural reading situations, and this might affect both the outcome and process measures applied. For example, the use of any kind of a head rest might improve the pupil size measurement, but this performance situation would be quite unusual for musicians. The same is true for the design where musicians have to read or judge single note probes. It is not only an uncommon reading situation for musicians, but it also does not allow to study the reading as a process. This study clearly demonstrates that many important adjustments in reading strategy occur before and after a sequence of interest, and these changes of the reading strategy stay unobserved in single note probe designs. However, a certain level of control for the study design is necessary: since music notation is a very complex symbol system, some control over the stimuli required or else it would be impossible to deduce what causes changes in the analyzed parameters.

To summarize the key findings of our study, incongruence in music notation not only increased the number of performance mistakes, but also led to changes in the reading processes on the eye-movement level. The processing difficulty increased for performances in the more complicated tonality. No significant interaction between incongruence and performance modality (singing or playing) as well as no significant relationship between cognitive characteristics and the number of correct performances of incongruent melodies were observed. The study of handling incongruences in music notation is important for music education, because there are many musical pieces built on these harmonic, melodic, or rhythmic surprising elements. In traditional

instrumental training, where the early stages of learning are typically built around tonal music, students begin to form expectations of the upcoming music—but when changing into atonal music, for example, these expectations do not hold anymore. In music pedagogy, it raises a question of how to prepare students to handle this kind of music. We propose the use of music-reading exercises with incongruent elements in them. In addition, the use of commercial eye-trackers might also be useful in music education, since they allow the visualization of the music-reading process to the performers themselves and provide a basis for discussions on what actually occurs during the reading. It also seems that the common belief of some that singing is the kind of music activity that everybody can easily perform in contrast to playing on a musical instrument where a special training is required is not correct. Our study demonstrates that when you have basic reading skills in both instruments, playing from notation on a piano is not as complicated as one might think.

### Acknowledgements

The authors thank Anna-Kaisa Ylitalo, Erkki Anto, and Johanna Kaakinen for their useful comments on the data analysis, the study participants for their participation and anonymous reviewers for their constructive comments and suggestions.

### Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The project was funded by the Academy of Finland (grant number: 278386) and the Finnish Cultural Foundation for the first author (grant number: 00180199).

### ORCID iDs

Natalia Chitalkina  <https://orcid.org/0000-0002-4083-8065>

Marjaana Puurtinen  <https://orcid.org/0000-0003-1972-2268>

### Supplemental material

Supplemental material for this article is available online.

### Notes

1. The statistical analysis of these parameters was presented at the ACM Symposium on Eye Tracking Research and Applications (ETRA 2019) and reported in the conference proceedings (Chitalkina, 2019).
2. We used the likelihood ratio test to compare the likelihood of a full and reduced model with and without the fixed effect under consideration. According to Wilks' theorem, negative two times the log likelihood ratio of two models approaches a chi-square distribution with degrees of freedom of the number of parameters that differ between the models (Winter, 2013).

### References

- Ahken, S., Comeau, G., Hebert, 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.

- Alexander, M. L., & Henry, M. L. (2012). The development of a string sight-reading pitch skill hierarchy. *Journal of Research in Music Education, 60*(2), 201–216.
- Arthur, P., Khuu, S., & Blom, D. (2016). Music sight-reading expertise, visually disrupted score and eye movements. *Journal of Eye Movement Research, 9*(7), 1–11.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. arXiv preprint arXiv: 1406.5823.
- Chitalkina, N. (2019). When you don't see what you expect: Incongruence in music and source code reading. In K. Krejtz & B. Sharif (Eds.), *Proceedings of the 11th ACM Symposium on Eye Tracking Research and Applications* (pp. 1–3). New York: Association for Computing Machinery.
- Chitalkina, N., Bednarik, R., Puurtinen, M., & Gruber, H. (2019). Prediction as a prerequisite of skilled reading: The cases of source-code and music notation. In N. Chitalkina, R. Bednarik, M. Puurtinen, & H. Gruber (Eds.), *Proceedings of the 19th Koli Calling International Conference on Computing Education Research* (pp. 1–9). <https://doi.org/10.1145/3364510.3364516>
- Fine, P., Berry, A., & Rosner, B. (2006). The effect of pattern recognition and tonal predictability on sight-singing ability. *Psychology of Music, 34*(4), 431–447.
- Gromko, J. E. (2004). Predictors of music sight-reading ability in high school wind players. *Journal of Research in Music Education, 52*(1), 6–15.
- Gunter, T. C., Schmidt, B.-H., & Besson, M. (2003). Let's face the music: A behavioral and electrophysiological exploration of score reading. *Psychophysiology, 40*(5), 742–751.
- Hadley, L. V., Sturt, P., Eerola, T., & Pickering, M. J. (2018). Incremental comprehension of pitch relationships in written music: Evidence from eye movements. *Quarterly Journal of Experimental Psychology, 71*(1), 211–219.
- Hayward, C. M., & Gromko, J. E. (2009). Relationships among music sight-reading and technical proficiency, spatial visualization, and aural discrimination. *Journal of Research in Music Education, 57*(1), 26–36.
- Herrero, L., & Carriedo, N. (2019). The contributions of updating in working memory sub-processes for sight-reading music beyond age and practice effects. *Frontiers in Psychology, 10*, Article 1080.
- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press.
- Huovinen, E., Ylitalo, A. K., & Puurtinen, M. (2018). Early attraction in temporally controlled sight reading of music. *Journal of Eye Movement Research, 11*(2). <https://bop.unibe.ch/JEMR/article/view/4094>
- Komogortsev, O. V., Gobert, D. V., Jayarathna, S., Koh, D. H., & Gowda, S. M. (2010). Standardization of automated analyses of oculomotor fixation and saccadic behaviors. *IEEE Transactions on Biomedical Engineering, 57*(11), 2635–2645.
- Kopiez, R., & Lee, J. (2006). Towards a dynamic model of skills involved in sight reading music. *Music Education Research, 8*(1), 97–120.
- Kopiez, R., & Lee, J. (2008). Towards a general model of skills involved in sight reading music. *Music Education Research, 10*(1), 41–62.
- Liao, H.-I., Yoneya, M., Kashino, M., & Furukawa, S. (2018). Pupillary dilation response reflects surprising moments in music. *Journal of Eye Movement Research, 11*(2), Article 13.
- Madell, J., & Hébert, S. (2008). Eye movements and music reading: Where do we look next? *Music Perception: An Interdisciplinary Journal, 26*(2), 157–170.
- McPherson, G. E. (1994). Factors and abilities influencing sightreading skill in music. *Journal of Research in Music Education, 42*(3), 217–231.
- Meinz, E. J., & Hambrick, D. Z. (2010). Deliberate practice is necessary but not sufficient to explain individual differences in piano sight-reading skill: The role of working memory capacity. *Psychological Science, 21*(7), 914–919.
- Nevala, J., Kairaluoma, L., Ahonen, T., Aro, M., & Holopainen, L. (2007). *Lukemis-ja kirjoittamistaitojen yksilötestistö nuorille ja aikuisille* [Assessment battery for reading disabilities in young and adults]. Niilo Mäki Institute.
- Olsen, A. (2012). The Tobii I-VT fixation filter. *Tobii Technology*. [http://www.vinis.co.kr/ivt\\_filter.pdf](http://www.vinis.co.kr/ivt_filter.pdf)

- Penttinen, M., Huovinen, E., & Ylitalo, A.-K. (2015). Reading ahead: Adult music students' eye movements in temporally controlled performances of a children's song. *International Journal of Music Education*, 33(1), 36–50.
- Puurtinen, M. (2018). Eye on music reading: A methodological review of studies from 1994 to 2017. *Journal of Eye Movement Research*, 11(2), 1–16.
- Reifinger, J. L., Jr. (2018). The relationship of pitch sight-singing skills with tonal discrimination, language reading skills, and academic ability in children. *Journal of Research in Music Education*, 66(1), 71–91.
- Sloboda, J. (1974). The eye-hand span—An approach to the study of sight reading. *Psychology of Music*, 2(2), 4–10.
- Sloboda, J. (1976). Visual perception of musical notation: Registering pitch symbols in memory. *Quarterly Journal of Experimental Psychology*, 28(1), 1–16.
- Tan, S.-L., Wakefield, E. M., & Jeffries, P. W. (2009). Musically untrained college students' interpretations of musical notation: Sound, silence, loudness, duration, and temporal order. *Psychology of Music*, 37(1), 5–24.
- Vuust, P., Dietz, M. J., Witek, M., & Kringelbach, M. L. (2018). Now you hear it: A predictive coding model for understanding rhythmic incongruity. *Annals of the New York Academy of Sciences*, 1423(1), 19–29.
- Weiss, M. W., Trehub, S. E., Schellenberg, E. G., & Habashi, P. (2016). Pupils dilate for vocal or familiar music. *Journal of Experimental Psychology: Human Perception and Performance*, 42(8), 1061–1065.
- Winter, B. (2013). Linear models and linear mixed effects models in R with linguistic applications. arXiv preprint arXiv: 1308.5499.
- Wolf, T. (1976). A cognitive model of musical sight-reading. *Journal of Psycholinguistic Research*, 5(2), 143–171