Eye movement patterns when playing from memory: Examining consistency across repeated performances and the relationship between eyes and audio

Lauren Fink

Dept. of Music, Max Planck Institute for Empirical Aesthetics, Frankfurt a.M., Germany
Dept. of Psychology, Neuroscience, & Behaviour, McMaster University, Hamilton, Ontario, Canada
finkl1@mcmaster.ca

While the eyes serve an obvious function in the context of music reading, their role during memorized music performance (i.e., when there is no score) is currently unknown. Given previous work showing relationships between eye movements and body movements and eye movements and memory retrieval, here I ask 1) whether eye movements become a stable aspect of the memorized music (motor) performance, and 2) whether the structure of the music is reflected in eye movement patterns. In this case study, three pianists chose two pieces to play from memory. They came into the lab on four different days, separated by at least 12hrs, and played their two pieces three times each. To answer 1), I compared dynamic time warping cost within vs. between pieces, and found significantly lower warping costs within piece, for both horizontal and vertical eye movement time series, providing a first proof-of-concept that eye movement patterns are conserved across repeated memorized music performances. To answer 2), I used the Matrix Profiles of the eye movement time series to automatically detect motifs (repeated patterns). By then analyzing participants' recorded audio at moments of detected ocular motifs, repeated sections of music could be identified (confirmed auditorily and with inspection of the extracted pitch and amplitude envelopes of the indexed audio snippets). Overall, the current methods provide a promising approach for future studies of music performance, enabling exploration of the relationship between body movements, eye movements, and musical processing.

Keywords: memory, piano, eye-tracking, dynamic time warping, matrix profile, motif, audio indexing

1. Introduction

Most eye-tracking research in musical contexts has focused on music (sight-)reading (Fink, Lange, Gröner, 2019; Madell, & Hébert, 2008; Puurtinen, 2018; Perra et al., 2022). While the eyes serve an obvious function in the context of music reading, their role during memorized or improvisatory music making contexts is currently unclear. What are the eyes doing when the performer(s) are not focused on a piece of paper?

Here, I explore memorized music performance – that is, music initially learned from a score but performed from entirely memory, with no visual aid of the score. In this context, there is nothing to look at per se. Performers can direct their gaze wherever they like. Perhaps they have consciously considered where to orient their gaze throughout a performance, or perhaps their eye movements have unknowingly become yet another highly rehearsed motor behavior accompanying their larger motor movements required to produce the music they are creating. A first step is, therefore, to ask whether there are any patterns in eye movement data during memorized music performance and, if so, how such patterns relate to the music being performed.

One can theorize that eye movements might be particularly interesting for a few different reasons: eye movements might reflect 1) cognitive processing of musical structures to-be performed; 2) memory-related processes; 3) motor-processes; 4) expertise; 5) specific acoustic events. For example, previous work has shown that eye movements and memory are intricately linked, with eye movements capable of reinstating spatiotemporal context and successful memory retrieval (Ryan & Shen, 2020). Further, eye movements can reflect mental schemas (Võ & Wolfe, 2013; Wynn et al., 2022) and additionally can be considered a planned, proactive part of bodily action sequences (Land, 2009).

Expert music performance involves years of developing fine motor skills. Such refinement seems to also involve ocular motor behavior, with many studies from the music reading literature showing different patterns of eye movement activity in experts vs. novices, such as increased parafoveal processing (Sheridan, Maturi, & Kleinsmith, 2020). Further, the eyes tend to look ahead of the what the hands are playing by 0.7-1.3 seconds, with the number of notes ahead increasing with expertise (novices: 2; experts 4+; Furneaux & Land, 1999; Perra et al., 2022). It is an open question whether experts and novices display different ocular motor behaviours when performing from memory. Before answering such a question, however, a general idea of what the eyes are doing when performing from memory is required.

Beyond memory, mental schemas, motor sequences, or expertise, eye movements might also be related to the produced audio. For example, previous work has shown that acoustic features, like increasing pitch, can alter ocular motor behavior (Schaefer et al., 1981; Xaio et al., 2007), or that complex audio, like movie soundtracks or narration, can influence visual exploration (Coutrout et al., 2012; Ross & Kowler, 2013). To date, there is very little study of which musical/acoustic features might drive eye movements in music listening or performance contexts.

Thus, in this proof-of-concept, pilot study, I ask: 1) whether eye movement patterns exhibit consistency across multiple performances of the same memorized music, and 2) whether the structure of the performed music is reflected in eye movement patterns?

2. Methods

2.1. Participants

Three pianist colleagues from the Max Planck Institute for Empirical Aesthetics volunteered for this pilot study. The experimental procedure was in accordance with guidelines ethically approved by the Ethics Council of the Max Planck Society.

2.2. Materials

Participants were asked to choose two pieces from their memorized repertoire, of 30 secs to 3 minutes in length (see Table 1 for chosen repertoire). They performed on the Steinway & Sons piano in the ArtLab at the Max Planck Institute for Empirical Aesthetics. Pupil Labs' Pupil Invisible glasses (Tonsen, Baumann, & Dierkes, 2020) were used to record eye movement data at 200 Hz. These glasses also recorded world video data at 30 Hz and world audio data at 48 kHz.

2.3. Procedure

On at least four different days, separated by at least 12 hours, participants performed each of their two pieces three times each. Before commencing performance, participants were outfitted with mobile eye-tracking glasses. The Pupil Invisible glasses do not require calibration, but offer the possibility for adjustment of the assumed eye model. Thus, calibration was checked by asking participants to gaze at areas in their near environment (e.g., piano stand, edge of keyboard). Following this procedure, a recording was started and participants placed the mobile device connected to the eye-tracking glasses on the piano bench, or in their pocket, whichever was most comfortable.

No score was used during the performance and no instruction was given to participants about anything regarding their gaze. Participants were instructed not to stop, even if a mistake was made. In some cases, the pianists were unhappy with their recordings and asked to make extra ones; in others, stoppages or technical errors resulted in an unusable trial. The pieces, number of trials, and average duration of each performance are presented in Table 1, by participant. In total, the data set consisted of 84 recordings.

2.4. Data Analysis

All data were processed in MATLAB (Mathworks, 2021b), using the Signal Processing, Statistics, Symbolic, and Audio System Toolboxes. Custom scripts additionally relied on the Matrix Profile code maintained here: https://matrixprofile.org/. Code required to produce the analyses below is available at: https://github.com/lkfink/icmpc2023.

For the audio analyses, amplitude envelope was extracted using the *envelope* function's 'rms' method, with a 50 ms sliding window. A third-order Butterworth filter was used to low-pass the extracted upper envelope at 50 Hz. To extract pitch from the recorded audio, the *pitch* function was used with the 'PEF: pitch estimation filter' method, 100 ms sliding windows, 75% overlap, a pitch range of 20:3300 Hz, and no smoothing.

Table 1. *Performance information by participant.*

S	Piece	# r	Dur. (secs)
1	*S'Wonderful*, Gershwin	16	72.09 (1.34)
	Sonata No. 12, Op. 26, excerpt from 1st mvmt, Beethoven	20	74.82 (2.30)
2	*Waltz in A Minor*, Chopin	12	52.93 (1.51)
	Bourrée from English Suite #2 in A Minor, BWV 807	10	30.20 (0.70)
3	*Prelude in C Major*, BWV 870	10	108.03 (4.18)
	Fugue in C Major, BWV 870	16	143.21 (3.44)

Note. S = subject; #r = num. of recordings. Dur = duration, Mean (std).

As is clear in Table 1, last column, performance duration varied across each recording. This potential issue was circumvented in two ways: 1) within participant, all data were trimmed to the shortest duration recording, across all recordings; 2) dynamic time-warping was used to assess similarity between eye movement time series to allow for more local variability.

3. Results

The first goal was to assess whether eye movement patterns were similar across repeated memorized performances of the same music. Matrices displaying the warping cost between all eye movement traces of each participant are shown in Figure 1, for both horizontal (left column) and vertical (right column) eye movements. The warping cost was significantly lower within vs. between pieces for both horizontal ($t(1184) = -16.71$, $p < .001$) and vertical eye movement data ($t(1184) = -2.54$, $p = .011$); see Figure 1, bottom row). Note that a lower warping cost is better (i.e., greater similarity in eye movements). The visible square patterns in each participants' data illustrate the potential robustness of this effect, already observable at the individual level. Smaller sub-squares within the larger piece-specific squares indicate similarities within recording day (recall that participants came to the lab on multiple days).

My second goal was to assess if eye movement data might relate to the musical structures being performed. There are a variety of ways to answer this question. Here, I wanted an automated approach that did not involve analyzing musical scores or extracting musical/audio features. Thus, I chose a "reverse" approach by first finding conserved patterns in eye movement data and then asking if those eye movement patterns indexed similar audio segments. I used a recent automated time series mining algorithm called the Matrix Profile (Zhu et al., 2018) to find “motifs” (in the time series sense, not the musical sense) in the horizontal eye movement data. I focused on horizontal eye movement data, given its larger

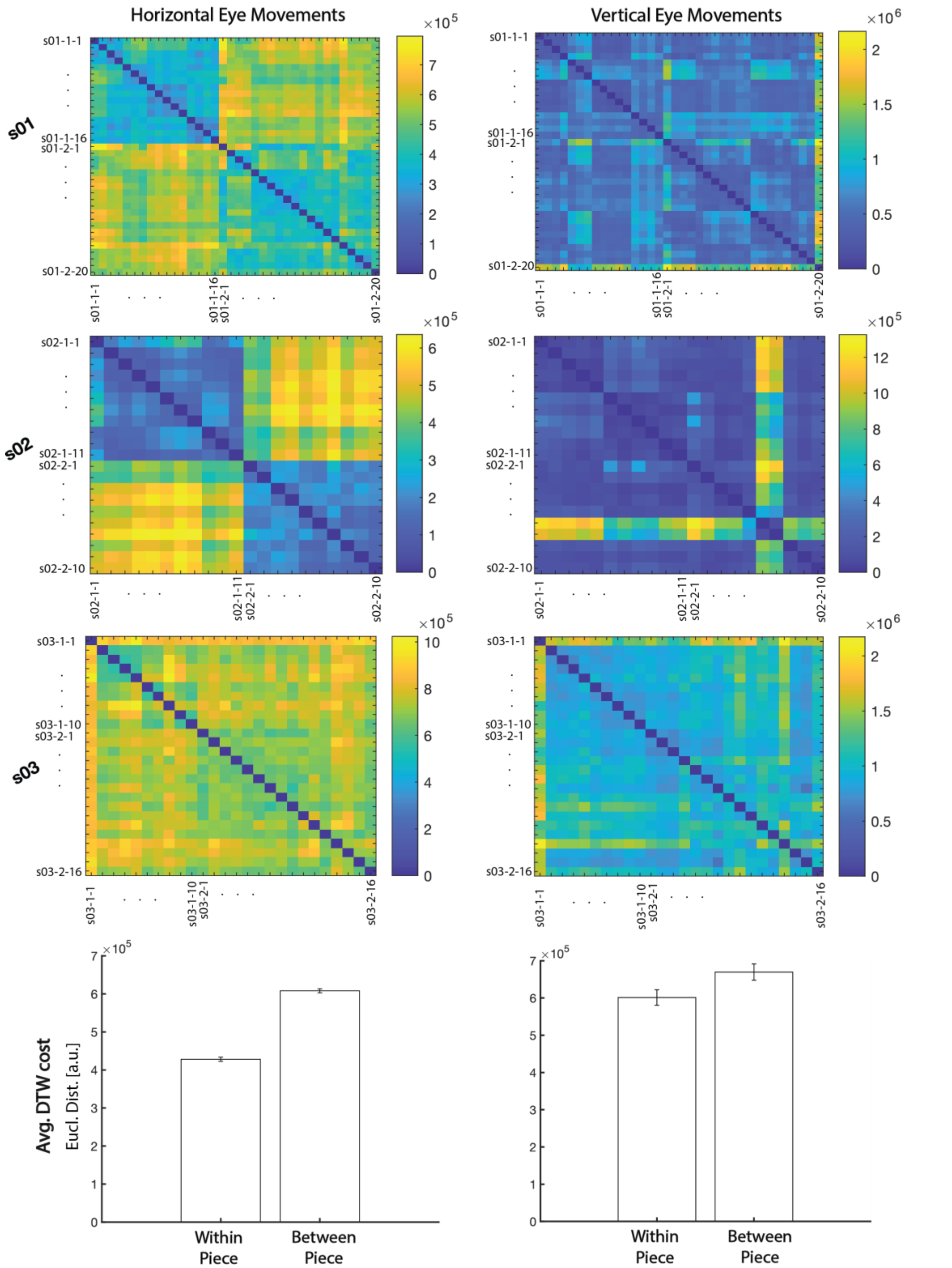

Figure 1. Warping cost between all horizontal (left column) and vertical (right column) eye movement traces of each participant (s01-s03), each recording (axis ticks). The average warping cost across all participants, within vs. between pieces, is shown in the bottom row. A lower value indicates less distance (greater similarity). Error bars represent SEM.

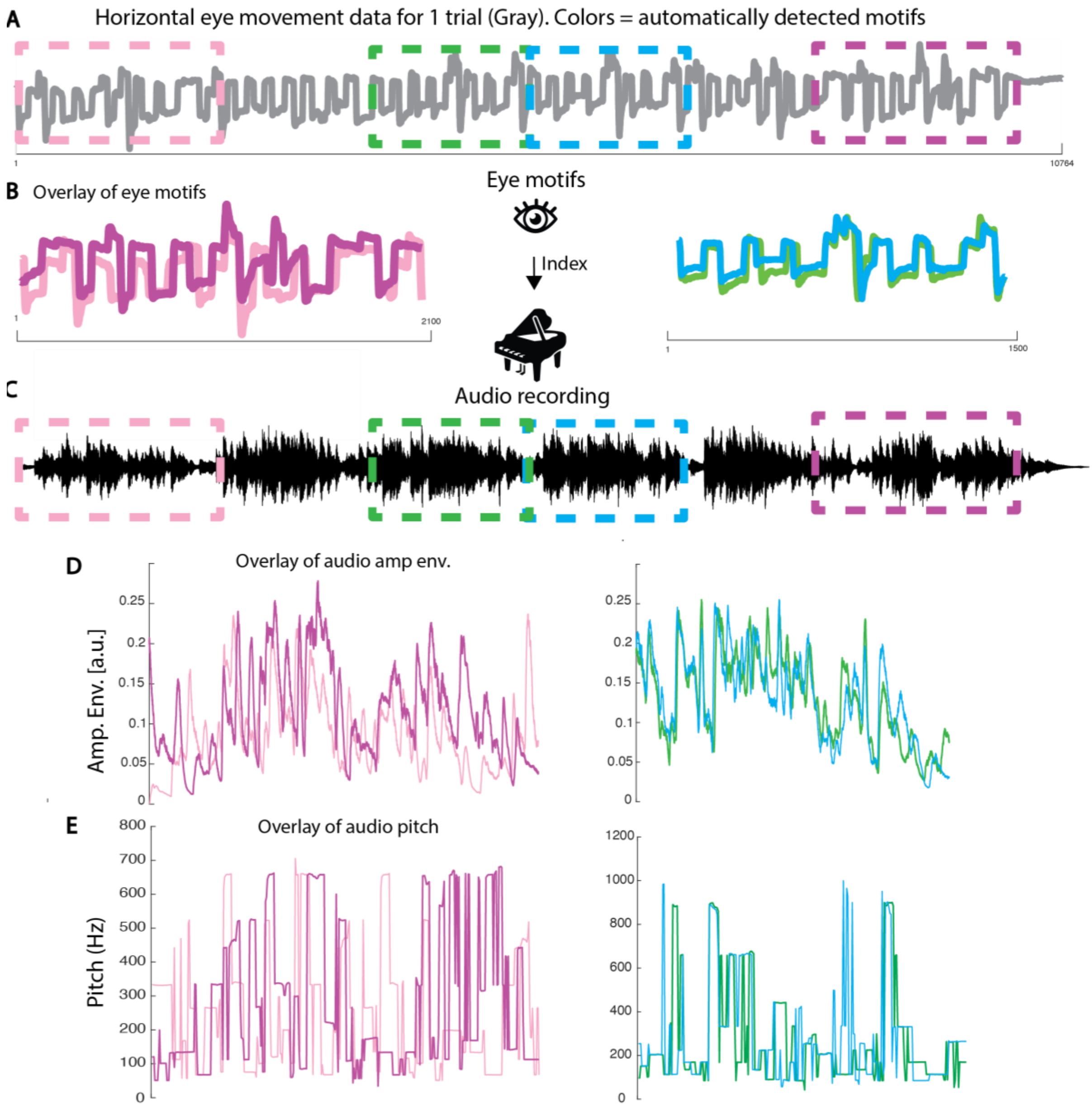

Figure 2. Single-trial processing pipeline to use automatically detected motifs in eye movement data for audio indexing. (A) Horizontal eye movement data from Participant 2, piece 1, recording 1. Pink/fuchsia and blue/green colored segments represent two detected motifs, respectively. Overlays of detected motifs are provided in (B). The temporal indices of those motifs were then used to index the audio recording of the performance (C). Both the amplitude envelope (D) and extracted pitch (E) are provided for audio corresponding to the respective eye movement motifs.

effect size in the previous analysis. I used the Pan Matrix Profile to confirm the appropriate time scales in which to search for motifs. In the example provided in Figure 2, two motifs of different lengths were found (see A-B, pink/fuchsia and blue/green segments, respectively). I then used the temporal indices of these motifs to index the audio from the recorded performance (Fig. 2C). Listening to the audio snippets is perhaps the best way to gain an intuition for the similarity of the indexed segments. Files can be found in the associated GitHub repository (https://github.com/lkfink/icmpc2023). In Figure 2D and E, the extracted amplitude envelopes and pitch time series from the indexed audio are plotted, respectively. As can be seen, there is a high degree of overlap in both of these time series. Though statistical testing is not appropriate at this stage, this exploratory analysis supports possibilities afforded by this technique for the future.

4. Discussion

This pilot study provides evidence that even when performers are not looking at a musical score, eye movements are a relevant measure of interest. By asking participants to perform the same pieces from memory on multiple occasions, it is possible to identify patterns of ocular activity across and within each performance. Similarity of eye movement trajectories within vs. between the two pieces each participant chose to play was confirmed via analyses showing significantly lower warping costs for both vertical and horizontal eye movements.

It is worth noting that vertical eye movement time series are fairly static, with variability driven largely by eyeblinks. Given the potential relevance but also relative sparsity of eyeblinks (Lange & Fink, 2023), future studies might use even longer memorized pieces to study ocular dynamics over greater time scales. Future work might also use the vertical eye movement time series to find structural breaks in the performed music, in line with previous research showing blinks can index auditory attention, often occurring at phrase or event boundaries and moments of lower cognitive load (Fink, 2014; Huber, Martini, & Sachse, 2022; Siegle, Ichikawa, & Steinhauer, 2008).

Horizontal eye data is generally more variable over time and, in the current case, showed a larger difference in warping cost, within vs. between pieces. It is possible that similarities in horizontal eye movement data could be driven by pianists looking at their hands. Future studies focusing on instrumentalists for whom viewing their hands and/or instrument is not possible/relevant would clarify this potential explanation and would allow for testing whether these effects translate to other instrumentalists. Additionally, studying the evolution of eye movement patterns throughout the learning / memory process might provide further insights into the mechanisms underlying similar patterning in eye movements during memorized performance. How do eye movements change from first reading to last reading to first memorized performance to last memorized performance? Do the eyes reveal the transition from learning to learned?

Other interesting future avenues of research include: 1) Comparing eye movements of performers who learn aurally vs. visually. What changes in ocular activity occur when transitioning from learning to memorized, in both contexts? How do eye movement patterns differ for the same piece learned in a different modality? Such studies allow for a better understanding of if/how visual representations of music affect ocular activity and learning.

2) Examining eye movement similarity during imagined vs. real performance. Such experiments would control for both the motor confound of performance, allowing to study if/how eye movements change when they are not embedded in larger motor sequences, and any potential confound related to looking at one’s hand(s) or instrument.

3) Exploring the relationship between dominant eye movement, body motion, and musical metric frequencies. Do these three temporal scales show systematic and consistent relationships with one another?

The automated motif-finding and audio indexing example further speaks to the connection between eye movements and musical structure. However, whether the similar eye movement patterns relate to similar music passages because of the cognitive processing of the music being performed or because of the motor act required to produce the music is a question for future research. To some degree the motor and cognitive aspects might be inherently inter-linked. Nonetheless, it is an interesting open question regarding which mental musical schemas, musical interpretations, produced motor acts, and/or produced acoustic features drive eye movement activity.

The current study already provides preliminary evidence for some sort of relation between eye movements and auditory material, with automatically detected eye movement motifs able to index similar music passages. This approach is particularly useful, given that it does not rely on lengthy music theoretic analyses of scores, or choosing which audio features are relevant and should be extracted computationally. Instead, time series mining algorithms can be applied to ocular data to find the temporal indices of similar patterns, which can then be used to find similar passages in the audio recording. In the future, using pre-labelled or segmented audio could be interesting for providing a better understanding of the types of musical passages that tend to be retrievable via eye movement patterns. Likewise, asking all participants to memorize the same piece would allow to test the robustness of this approach.

Ultimately, developing the methodological approach proposed in this pilot study – which unites techniques from music performance analysis, music information retrieval, time series mining, and psychophysiology – will hopefully contribute to a deeper and more holistic understanding of the whole-body motor act of music performance and the cognitive strategies underlying its execution.

References

Coutrot, A., Guyader, N., Ionescu, G., & Caplier, A. (2012). Influence of soundtrack on eye movements during video exploration. *Journal of Eye Movement Research*, *5*(4), 2.

Fink, L. (2014). Music modulates eyeblinks: An examination of temporal coordination (Masters thesis, University of Cambridge).

Fink, L. K., Lange, E. B., & Groner, R. (2018). The application of eye-tracking in music research. *Journal of Eye Movement Research*, *11*(2).

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. Series B: Biological Sciences, 266(1436), 2435-2440.

Huber, S. E., Martini, M., & Sachse, P. (2022). Patterns of eye blinks are modulated by auditory input in humans. *Cognition*, *221*, 104982.

Lange, E., & Fink, L. (2023, in press). Eye-blinking, musical processing, and subjective states—A methods account. *Psychophysiology*, e14350.

Madell, J., & Hébert, S. (2008). Eye movements and music reading: Where do we look next? *Music Perception*, *26*, 157–170.

Perra, J., Latimier, A., Poulin-Charronnat, B., Baccino, T., & Drai-Zerbib, V. (2022). A meta-analysis on the effect of expertise on eye movements during music reading. *Journal of Eye Movement Research, 15*(4).

Puurtinen, M. (2018). Eye on music reading: A methodological review of studies from 1994 to 2017. *Journal of Eye Movement Research, 11*(2).

Ross, N. M., & Kowler, E. (2013). Eye movements while viewing narrated, captioned, and silent videos. *Journal of vision*, *13*(4), 1-1.

Ryan, J. D., & Shen, K. (2020). The eyes are a window into memory. *Current Opinion in Behavioral Sciences*, *32*, 1-6.

Schaefer, K. P., Süss, K. J., & Fiebig, E. (1981). Acoustic-induced eye movements. *Annals of the New York Academy of Sciences*, *374*, 674-688.

Siegle, G. J., Ichikawa, N., & Steinhauer, S. (2008). Blink before and after you think: Blinks occur prior to and following cognitive load indexed by pupillary responses. *Psychophysiology, 45*(5), 679-687.

Sheridan, H., Maturi, K., & Kleinsmith, A. L. (2020). Chapter Five - Eye movements during music reading: Toward a unified understanding of visual expertise. *Psychology of Learning and Motivation, 73*, 119–156.

Tonsen, M., Baumann, C. K., & Dierkes, K. (2020). A high-level description and performance evaluation of pupil invisible. *arXiv preprint arXiv:2009.00508*.

Võ, M. L. H., & Wolfe, J. M. (2013). The interplay of episodic and semantic memory in guiding repeated search in scenes. *Cognition*, *126*(2), 198-212.

Wynn, J. S., Van Genugten, R. D., Sheldon, S., & Schacter, D. L. (2022). Schema-related eye movements support episodic simulation. *Consciousness and Cognition*, *100*, 103302.

Xiao, M., Wong, M., Umali, M., & Pomplun, M. (2007). Using Eye-Tracking to Study Audio—Visual Perceptual Integration. *Perception*, *36*(9), 1391-1395.

Zhu, Y., Yeh, C. C. M., Zimmerman, Z., Kamgar, K., & Keogh, E. (2018, November). Matrix profile XI: SCRIMP++: time series motif discovery at interactive speeds. In *2018 IEEE International Conference on Data Mining (ICDM)* (pp. 837-846). IEEE.