Spring 2022

Eye-movements of Vocal Performers Across Experience Levels

Charlotte Kelly
cmkelly@rollins.edu

Eye-movements of Vocal Performers Across Experience Levels

Charlotte Kelly

Senior Honors Thesis

Department of Psychology, Rollins College

Abstract

Expertise, such as music expertise, is commonly studied through an analysis of eye-movements. Experts typically have fewer fixations, longer saccade amplitudes, and thus greater perceptual spans when reading music than non-experts. Most musical expertise literature is focused on instrumentalists and sight-reading. The current study aimed to extend the research to include vocalists and to see if there are still expertise effects when both experts and non-experts are familiar with the piece of music. Participants were recruited to sing a piece from their choir once when they had first started learning the piece and again right before their concert. They were separated into three expertise groups based on their responses to the Vocal Experience Questionnaire. There was a significant interaction between test type (pre-test/post-test) and experience level (Low/Medium/High) on fixations. The Low Level group had a marginally significant decrease in fixations between sessions. There were no other significant main effects or interactions. These results are contrary to the literature, and thus there are some methodological adjustments that could be made in the future to get more representative results, including using a different piece of music, adjusting the Vocal Experience Questionnaire questions, and increasing the participant size. Through the marginally significant decrease in fixations for the low experience group, one can conclude that practice helps close the gap between non-experts and experts. Ways this research could be expanded include focusing on the “mastery” versus “expertise” hypothesis, conducting a psychophysics experiment instead of a cognitive one, and including eye-voice span in the analysis.

Eye-movements of Vocal Performers Across Experience Levels

The study of expertise is a growing topic in cognitive psychology, and researchers keep finding more ways to differentiate experts from amateurs and novices. This is especially the case with eye-movements, in which scientists have found that experts tend to look at their area of expertise in a different way than that of non-experts and allows them to better process the visual information relevant to that area.

For instance, Papesh et al. (2021) studied this phenomenon using a visual search task, which is a task where participants are asked to search on an image known as a visual array for a set of target items that may or may not be there. Not only were they interested in seeing an expertise difference, they were also interested in seeing how the accumulation of expertise might affect eye-movement patterns. To do this, the researchers amassed ten naïve participants and taught them to complete the visual search task across fourteen sessions, recording their eye-movements each session using an eye-tracker so that they could see the expertise change. They found that as participant accuracy increased (and thus they were becoming experts at the visual search task), the number of times their eyes visited individual items in the visual array decreased. These visits are known as *fixations*. Additionally, they found that the participant's glances across the visual array were longer as expertise grew. These glances between fixation points are known as *saccades*, and the longer the glance is, the longer the *saccade amplitude*. These two eye-movement patterns are important because they show that the participants are gaining a greater *perceptual span* as they become experts. This means that the experts are looking at the scene holistically and can take in more information in a single glance that non-experts can, a skill that is beneficial in any sort of visual task, especially in visual search. In summary, the findings of Papesh et al. (2021) show that novices tend to dart their eyes around visual arrays, look at many objects, and might take a while to find what they are looking for, while experts look at the visual array with larger, sweeping glances and take in all of the information they need before fixating on the target object. This is a very important finding, for it not only shows the eye-movement growth of participants as they progressed from novice to expert, but it also acts as a wonderful culmination of the past fifty years of visually-based expertise research.

The phenomenon described by Papesh et al. (2021) goes beyond basic visual search tasks and has been observed by researchers in all kinds of visually guided expertise. For instance, expert chess players have fewer fixations and larger perceptual spans when viewing the chess board than their novice counterparts (Reingold et al., 2001), and even expert radiologists follow a similar pattern when searching lung x-rays for cancer nodules (Drew et al., 2013). Music is another area of expertise where these findings apply, but it is relatively understudied compared to other areas of expertise such as text reading (Madell & Hérbert, 2008). Various studies have examined musical expertise through eye-movements, but the list is short and narrow; however, the basic research so far does agree with other expertise studies, showing the difference between the fixations, saccades, and perceptual spans of novices and experts during the music reading process.

Patterns of Fixations in Music Reading

Fixations are when the eye focuses on a specific spot on a visual stimulus. In studies of musical expertise, researchers study these fixations in reference to reading written sheet music. Drai-Zerbib and Baccino (2018) asked expert and non-expert musicians to listen to and read various excerpts of music to determine discrepancies between them. The experts were more accurate in this task than the non-experts, and eye-movement data revealed experts to have fewer fixations and shorter fixation durations compared to the non-experts. Through a single fixation, the experts could extract more information than their non-expert counterparts, and thus did not need to make as many fixations overall. This partially replicates the findings of Papesh et al. (2021) and shows the universality of eye-movement data across different types of expertise.

Penttinen and Huovinen (2011) found similar results in their study on sight-reading. Sight-reading is when someone reads a piece of music that they have not seen before. If they then subsequently perform this piece, they are described as sight-playing or sight-singing it. In their study, Penttinen and Huovinen (2011) taught sight-reading skills to a group of fifteen novice musicians so that they could study how expertise forms. They asked the participants to sight-play various excerpts on the piano while their eye-movements were recorded by an eye-tracker. The researchers found that, although fixations did decrease as expertise increased, the *pattern* of fixations also changed. When the novices first began the training, the number of fixations tended to increase as they continued through the piece. Near the end of their training, there was only a larger proportion of fixations at the end of each *measure*, which is the equivalent of a “sentence” of music. This shows that as the participants gained musical expertise, they gained a greater understanding of the metrical divisions inherent in music. Written music is separated in this way for ease of reading, but those who are not experts in reading music would likely ignore measure lines completely, as is represented by this data.

Patterns of Saccades in Music Reading

Saccades are the glances between individual fixations, and saccade amplitudes are the length of that glance. For reading music, this may be the distance between individual notes or note clusters that a participant fixates on. Penttinen et al. (2013) continued their musical expertise research from before, this time extending their investigation into saccade amplitudes as well. They separated participants into three groups of varying expertise (novice, lower-level amateur, higher-level amateur) and asked them to silently read various music excerpts and then describe them. Once again, the eye-tracking data matched that of the previous studies; the higher the level of expertise, the fewer fixations, shorter durations, and longer saccade amplitudes. One additional correlation they found was that participants with higher expertise also participated in more linear scanning. This means that instead of the participants' eyes jumping around on the page, they would follow each line of music linearly. This important piece of evidence shows that there are some ways of assessing eye-movements in experts that may vary across different types of expertise. For instance, in the Papesh et al. (2021) study, they cannot assess for "linear scanning" because objects are not ordered linearly in a visual search task. In music tasks or any reading task, however, linear eye-movements may be especially important for experts to find target items and stay in time with the music.

These findings also apply to performing sight-read music, as verified by Truitt et al. (1997). They had pianists in two experience groups (skilled and less skilled) perform sight-read music while their eye-movements were recorded and also found that the more expert pianists had shorter fixation durations and longer saccade amplitudes.

Perceptual Span

All of this data culminates into an understanding of perceptual span and holistic view. As discussed in Papesh et al. (2021), fewer fixations and longer saccade amplitudes are evidence for a view of space that is global, holistic, and "big picture," rather than focusing on individual details. Wong and Gauthier (2010) explicitly studied this perceptual span in experts. They took participants of various music reading levels and presented them with two almost identical note-clusters in quick succession, either in the center of their view or in the periphery. The second set of note-clusters in each trial contained arrows pointing to a singular note, and the participants would have to identify whether that note was altered from the first presentation. Their results showed that all of the participants, novices and experts, exhibited some level of a holistic view based on the speed and accuracy of their responses, but that this was more automatic in experts, showing that the expert musicians indeed viewed the musical stimuli with a more global view than their counterparts.

Burman and Booth (2009) did a similar experiment, but like Penttinen and Huovinen (2011), were more interested in the process of becoming an expert. They tested participants from two groups (skilled and less skilled) on a task in which they had to determine whether a “variant note” appeared in a melodic sequence (Burman & Booth, 2009). In the beginning, there was a significant difference in the perceptual spans of the two groups, with the skilled group performing better, but by the end of 20 rehearsals, the difference in perceptual span had diminished and was no longer significant. These findings show that practice of a specific musical task can build expertise in less experienced groups, allowing them to adopt a similar perceptual span to their experienced counterparts and further developing their visual processing of information to approach that of experts.

Current Study

In all of the studies above, there are a few common methods. Each asks their participants to either read and identify written music or to perform on an instrument, most commonly on the piano. Additionally, they investigate this expertise difference using sight-read pieces and excerpts. In the present study, I attempt to fill a gap in the literature and identify how these expertise differences may apply to *vocally* performed pieces; furthermore, I attempt to clarify the difference between expertise in a domain, such as music, versus “mastery” of a task, such as a specific piece of music. If all expertise groups have mastery of a piece of music, will there still be a difference in visual processing, or will practice make the eye-movements of the novices equivalent to that of the experts?

If participants' eye-movements are recorded while they sing a piece of music mostly unfamiliar to them, I expect there to be fewer fixations and longer saccade amplitudes for participants with greater domain-specific musical expertise, thus indicating that they have a greater perceptual span than their non-expert counterparts (like in Penttinen et al. (2013)). When participants' eye-movements are recorded in a second session after they have achieved a relative "mastery" of the piece, I expect all participants' fixations regardless of expertise to decrease and saccade amplitudes to increase (like in Burman and Booth (2009)). The more expertise a participant has, the smaller the margin of improvement should be and vice versa; however, contrary to Burman and Booth (2009), I do not believe that the results of all participants will converge in the final session, rather I believe there will still be a significant difference between participants of varying expertise levels. If this final hypothesis is correct, then it will provide evidence for the difference between "mastery" of a piece and "expertise" in a domain. If it is incorrect, then it will instead verify the common trope that "practice makes perfect," or rather "practice makes expert," for the non-expert participants' eye-movements will match those of the experts, despite them not being "experts" themselves.

Method

Participants

Twenty-two students participated in the study. All participants were members of the Rollins College Choir and were recruited in person during their choir rehearsals. Participants had 9.82 years of choral experience on average (*SD* = 4.98), ranging from 0 to 18 years. There were 10 Sopranos, 4 Altos, 3 Tenors, and 5 Basses. Participants were paid $10.00 in the form of an

Amazon gift card at the end of the study, given their successful and complete participation in both sessions. Participants were not excluded from participation based on race or age, with the only requirements being that participants must be at least 18 years of age, have normal or normal-to-corrected vision (e.g., eye-glasses), and be a regular member of the Rollins Choir.

Experimental Design

This study followed a 3 (Experience Level: Low Level, Medium Level, High Level) x 2 (Test Type: Pre-test, Post-test) mixed factorial design, with Experience Level as the between-groups variable.

Apparatus and Materials

I assessed my participants using a single laptop computer Lenovo ThinkPad P14 with a screen resolution of 1920 x 1080 and a screen length of 14 inches diagonally. Additionally, I recorded participants' voices using a small USB microphone, and recorded their eye-movements using a Tobii Pro Nano eye-tracker (Tobii Pro AB, 2014). Participants completed a single task twice across two sessions in which they sang the piece *The Virgin's Cradle Hymn* while reading the PDF version of the written sheet music (Rubbra, 1926). The participants also sang along with an audio recording of a choir singing the piece (Murray, 2019) in order to best replicate the choral experience. I chose the choral piece in collaboration with the director of Rollins Choir as a piece that the choir would sing at their end of semester concert. I used the Tobii Pro Lab software to manage all experiment and data recording procedures (Tobii Pro AB, 2014).

Reading the music on the computer screen did not affect the analysis, as the literature shows that there is no difference in performance or processing for those who read from printed music versus electronic music (Picking, 1997). Additionally, by having the participants read and sing a shared musical piece instead of different pieces, the experiment accounts for and limits the amount of natural variation of fixations and saccade amplitudes that would be present in pieces with different tempos and notations (Kinsler & Carpenter, 1995). By choosing this piece in coordination with the natural learning structure of the choir, I was able to test participants once before they started rehearsing and once again before their final performance of it, at which time they would have an expected “mastery” of the piece.

The sheet music stimulus was presented in slides to the participants that contained one system, or line, at a time (see Figure 1). Participants could move on to the next system by pressing the [SPACE] key. A total of five systems were presented (see Appendix A for a complete copy of *The Virgin’s Cradle Hymn*). Participants read the Latin text, despite the English translation being below.

Figure 1.

Second System of The Virgin’s Cradle Hymn

Measure

I created the Vocal Experience Questionnaire to record the relevant demographic information needed to split participants into experience groups. It contained questions regarding how long the participant had been in choir, how long they have been taking voice lessons, their class year, and their voice lesson "level" if they were taking the Applied Music course at Rollins College. I also included additional questions regarding experience with *The Virgin's Cradle Hymn* specifically, such as if they had seen the piece before and how often they rehearsed it. The question types included multiple choice and slide bar values (such as for number of minutes rehearsed). See the appendix for the full list of questions and answer choices (Appendix B).

Procedure

Once on-site, participants were given a consent form highlighting the voluntary nature of the study, their ability to withdraw from the study without penalty (other than the penalty of withheld compensation), the estimated length of the study (i.e., 10-15 minutes), and the general procedure of the study. Participants were given time to read the form, ask questions, and sign confirming their consent and age before the experiment began. Participants were informed that I would be recording their eye-movements using the video-based eye-tracker and that the recordings would not present any risk to participant vision. Their likeness was not recorded and participants were ascribed a random subject number, hence the data collection remained anonymous. Participants were also informed that I would be recording their voices using the microphone, but that their voices would not be analyzed in terms of performance, and were instead being recorded in order to record timing between the music and their eye-movements. I also emphasized that participation in this study would in no way affect their grade in choir and that their recorded voices would not be shared, furthering anonymity.

After a standard nine-point eye-tracking calibration, I presented the participants with the sheet music stimuli, "counted in" the choral track, and then the participants read the music and sang along with the recording as though they were performing the piece in front of an audience.

Participants were able to “flip pages” in their sheet music by pressing the [SPACE] key on the computer keyboard. Due to the sensitive timing of the eye-tracking recording, participants were not allowed to go back to previous pages. The eye-tracking recording automatically stopped after participants pressed the [SPACE] key on the fifth and last page.

Once they had finished the testing portion of the experiment, I redirected the participants to Qualtrics, an online survey website, to complete the Vocal Experience Questionnaire. I then prompted the participants to voluntarily schedule a second testing session. We repeated the above procedure at the following session, including the Vocal Experience Questionnaire in order to record the self-reported changes in practice. I conducted the second session with each participant an average of four weeks following their first session and then provided them with their $10.00 compensation following completion of the second testing session.

Results

I removed one participant from the analysis due to a recording failure and a second participant due to inappropriate behavior. Thus, I analyzed a total of 20 participants in the final analysis. Additionally, I removed two systems from the analysis of one participant because they accidentally skipped one system during experimentation and thus spent an unrepresentative amount of time on the following system.

Based on the responses of the participants, I chose “voice lesson level” as the expertise measure for this analysis. I did not choose “Class year” as the expertise measure because I felt it did not properly represent musical experience, nor did I choose “years in choir” as the measure because of the uneven variability across participants. In contrast, “voice lesson level” properly represented expertise, as students cannot move to the next level until they have reached specific criterion determined by the music department. The different “voice lesson levels” in the music major are 201C, which is the entrance level for voice in the music major; 201P, which is the "performance level" and is where vocal students spend the majority of their time; 301, which is the advanced level; and 401, which is by invitation only. In order to promote from 201P to 301, students must submit a petition to promote and be held under the same standards as 301 students for a full semester, showing their ability to be "more advanced." Using these voice lesson levels, I was able to easily group participants into three, mostly even experimental groups: the Low Level experience group (no voice lessons or 201C) with six participants, the Medium Level experience group (201P) with seven participants, and the High Level experience group (promoting from 201P-301, 301, and 401) with seven participants.

I analyzed all of the data in JASP statistical software (JASP Team, 2020), set the alpha level for the analyses at .05, and subjected the degrees of freedom for all significant comparisons to Tukey HSD corrections. The primary dependent variables were number of fixations and length of saccade amplitudes. I utilized two separate 3 (Experience Level: Low Level, Medium Level, High Level) x 2 (Test Type: Pre-test, Post-test) mixed model analyses of variance (ANOVAs), with experience level being the between-groups variable across both primary measures.

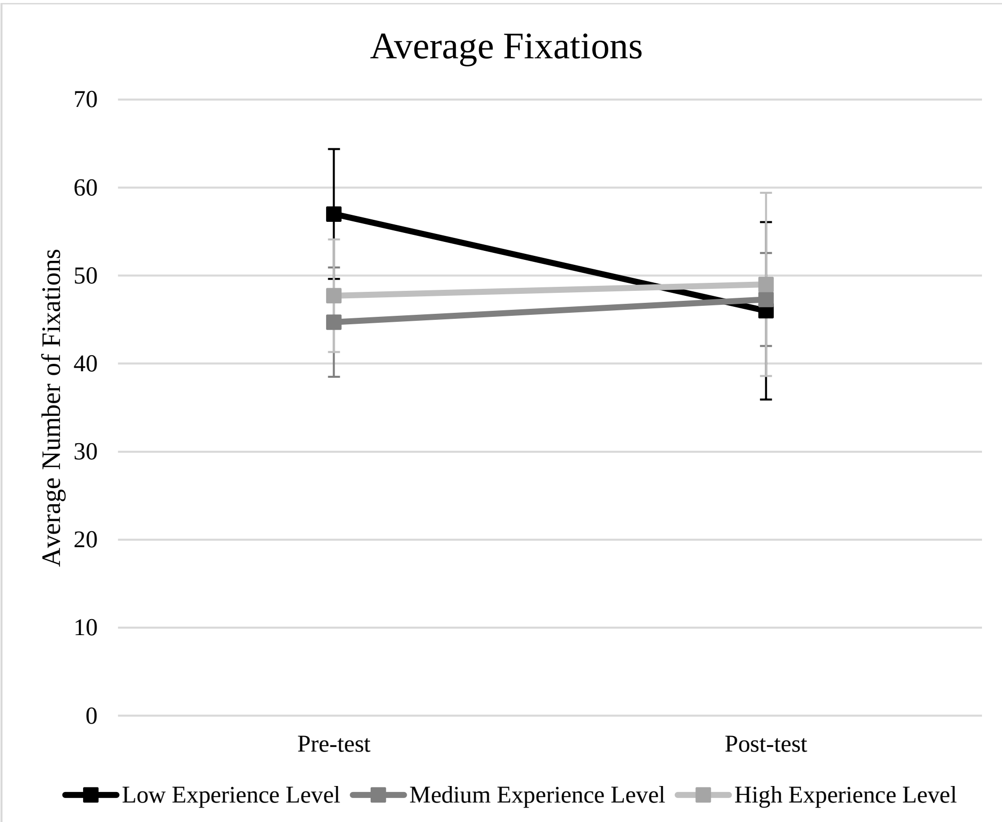
Average Fixations

The eye-tracking program recorded the average number of fixations per system for each expertise group. These recorded fixations included any time that the eyes rested on the screen, thus including both item fixations (e.g. on notes, text, dynamic markings) and fixations in blank space. There was a significant interaction between experience level and test type on fixations $F(2,17) = 4.74$, $p = .023$, $\eta^2 = .132$. The post-hoc analysis showed that those in the Low Level group had a marginally significant decrease in fixations between the pre-test ($M = 57$, $SD = 7.376$) and the post-test ($M = 46$, $SD = 10.08$) conditions, $p = .058$. The Medium Level and High

Level groups did not show a significant change between pre-test and post-test (see Figure 2), nor were there any significant main effects of test type or experience level on fixations, *p*s > .05.

Figure 2.

Average Number of Fixations per System



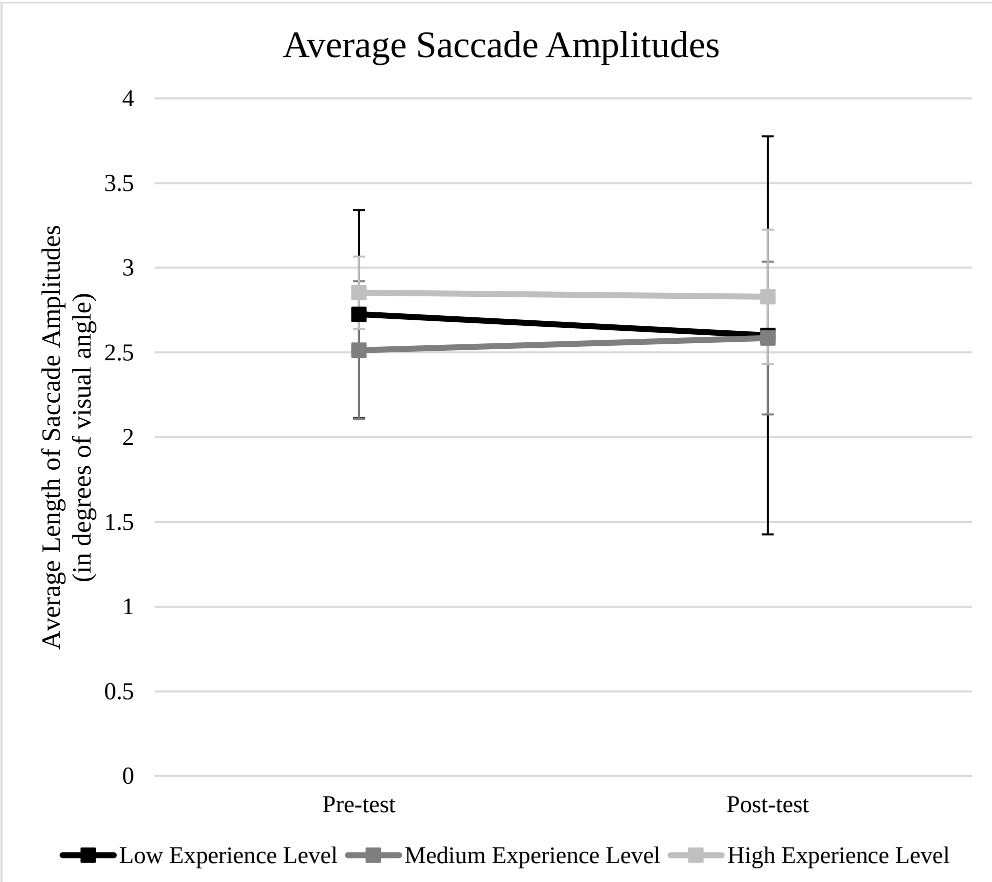
Note. The error bars represent the standard deviations for each data set and are color coded to match the line they correspond to for clarity.

Saccade Amplitudes

The eye-tracking program also recorded the average length of saccade amplitudes, or movements of the eyes from one fixation point to another, per system for each expertise group. There was not a significant interaction between experience level and test type on saccade amplitudes $F(2,17) = .196$, $p = .824$, $\eta^2 = .005$. There were also no main effects of test type or experience level on saccade amplitudes, *p*s > .05 (see Figure 3).

Figure 3.

Average Amplitude of Saccades per System



Note. The error bars represent the standard deviations for each data set and are color coded to match the line they correspond to for clarity.

Discussion

For this study, I hypothesized that for the pre-test, the group with the most experience (High Level) would have the fewest fixations and the longest saccade amplitudes, the group with the least experience (Low Level) would have the most fixations and the shortest saccade amplitudes, and the Middle Level group would be somewhere in the middle. In the post-test, I hypothesized that the number of fixations would decrease and the length of saccade amplitudes would increase for all three groups, with the Low Level group improving the most, thus closing the gap between them and the High Level group without reducing it completely.

The data supported some portions of my hypotheses, but most of my hypotheses were not supported. The different participant groups did not have significantly different fixations or saccade amplitudes, regardless of the session, nor did the participant groups improve across sessions. The data also did not trend in the direction of the most experienced (High Level) group having the least fixations and the longest saccade amplitudes, and thus the largest perceptual span. This is contrary to the findings of Drai-Zerbib and Baccino (2018), Penttinen et al (2013), Truitt et al. (1997), and Wong and Gauthier (2010). There are many methodological elements in this study that researchers could improve on in future studies to reduce the inconsistency between the current data and the data of the previous literature.

The Virgin's Cradle Hymn was chosen as the primary stimulus because of the piece's simplicity, but this may have been a draw back as well. Given the piece was slow and straightforward, it may not have allowed for much "growth" between sessions (i.e., a ceiling effect), and thus led to the lack of significance. Additionally, the director of the choir informed me that he no longer planned on including the piece in their final concert, and thus the practice time per participant between the first session and the second session was quite low ($M_{minutes\ per\ week} = 3.05$, $SD = 4.10$). It is questionable whether the participants actually achieved "mastery" of the piece before the second session, especially since they would have been expected to know many other songs for their concert instead, and thus would have been focusing their practice elsewhere.

This experiment also did not account for all forms of musical expertise. Although there were questions regarding vocal experience in the questionnaire, there were not any questions about instrumental experience, which could have very well influenced the performance of some participants in these tests. For instance, a participant could have been in the Low Level group because they are not taking voice lessons, but they might have a great understanding of music as a 301 level piano student. Adjusting the Vocal Experience Questionnaire to include questions that address other areas of musical expertise would eliminate these confounding variables. Finally, this was a relatively small sample size for a cognitive study. In the future, recruiting just a few more students (perhaps 10 in each group) could be the difference between marginal significance and significance for the Low Level group.

The one significant interaction between experience level and test type did partially support my second hypothesis. The Low Level group did marginally improve across the pre-test and post-test sessions by decreasing in the number of fixations and thus approaching the fixation count of the High Level group. This supports the findings of Burman and Booth (2009) and Penttinen and Huovinen (2011) and shows how practice can allow a non-expert to replicate the results of an expert. Sadly, there is not a distinction between the Low Level and High Level fixations in the post-test, so an argument cannot be made to distinguish “mastery” in a single piece and “expertise” in a domain at this time, but there are various ways that this experiment could be expanded to better test this theory.

Future Directions

In order to better distinguish “mastery” of a piece from “expertise” in a domain, researchers could replicate this experiment with an additional session that focuses on an unrelated music piece. For instance, participants could do the pre-test with a particular piece, a post-test with that same piece, and then a post-post-test with the unrelated piece. This way, researchers could identify whether the skills learned in the “mastery” of the primary piece translated over to the processing of the secondary piece, thus indicating a generalized and domain specific “expertise.”

One of the limitations outlined above for this experiment was the sample size. Given this study is cognitive in nature, this critique is appropriate; however, this line of research may benefit from a longitudinal approach in contrast to a population sampling approach. If this experiment were replicated longitudinally, the researchers would instead study fewer participants but over more sessions and over a longer period of time. This way, the researchers would gain a firm understanding of how these particular individuals learn, which could then be applied to a larger population. This alternative methodology could be a beneficial way to study this phenomenon in addition to the population sampling methods.

Finally, one measure of musical expertise that was not included in this paper, but certainly should be in future studies, is the eye-hand span. This is one of the measures that Truitt et al. (1997) utilized in addition to perceptual span and saccade amplitudes to analyze expertise differences. This measure describes the difference between what a musician is looking at in the sheet music and what the musician is currently playing. The practice of “looking ahead in the music” has always been a favored one of music instructors, and Truitt et al. (1997) verified that it is something that experts in the field tend to do. They looked specifically at the number of notes between when the participants looked at, or fixated on, a note and when they played that note on the piano. They found that experts had a larger eye-hand span than non-experts, meaning they were looking a significant number of beats ahead of what they were playing compared to the non-experts.

Furneaux and Land (1999), however, show that the eye-hand span can also be determined by recording the amount of time between where the eyes are looking and what the hand is playing. In their study, they defined this way of measuring the eye-hand span as the “time index,” defined the note-interval way of measuring the eye-hand span as the “note-index,” and recorded both indexes in their experiment with amateur and professional pianists. They asked them to sight-read and sight-play pieces of various complexity and found that the experts had a greater note index than the amateurs. There was not, however, a significant difference between the time indexes of the experts and amateurs. Given every participant sight-read a different piece, the researchers were unable to successfully determine why the time index was not significant, but the note index was. They theorized that either the experts’ pieces had naturally faster tempos and/or that the amateurs had more hesitations and mistakes while playing that slowed them down comparatively.

According to Rosemann et al. (2016), the eye-hand span is also relatively unaffected by practice. In their study, they recorded the eye-movements of nine pianists while they sight-played accompaniment for a short flute melody. They then gave the participants time to practice before repeating the test three more times and found that the eye-hand span only slightly changed across the four trials; instead, changes in tempo and complexity of the piece had a greater effect on the eye-hand span. In their discussion, they theorize that significant changes in eye-hand span would only be affected by long term practice. Thus, it is possible that the eye-hand span could be a strong component of any studies where researchers want to focus on the practice element of expertise and the potential difference between mastery and expertise.

Researchers have mostly studied the eye-hand span through concert pianists, but have more recently expanded their research to include violinists (Wurtz et al., 2009) and even xylophone players (Marandola, 2019). In the case of the xylophone players in Marandola (2019), the eye-hand span was renamed the eye-stroke span in order to better represent the performance style of this instrument class. When this measure is expanded to vocalists, it should appropriately be renamed the eye-voice span. Given the small scope of this project, the eye-voice span was not included in the analysis despite the collection of auditory data. For my next project in particular, it is my goal to analyze the eye-voice span in conjunction with fixations and saccades to differentiate the eye-movements of experts and non-experts.

As is often the case with science, this experiment has given us more questions than answers and has left the door open for more experimentation into the topic of musical expertise. With the significant interaction and the improvement of the Low Level group present in this study, it is clear that the answers are there and that simply an adjusted methodology must be used in order to find them.

References

Burman, D. D., & Booth, J. R. (2009). Music rehearsal increases the perceptual span for notation. *Music Perception*, *26*(4), 303-320. doi.org/10.1525/mp.2009.26.4.303

Drai-Zerbib, V., & Baccino, T. (2018). Cross-modal music integration in expert memory: Evidence from eye-movements. *Journal of Eye Movement Research*, *11*(2):4. doi.org/10.16910/jemr.11.2.4

Drew, T., Vo, M. L.-H., Olwal, A., Jacobson, F., Seltzer, S. E., & Wolfe, J. M. (2013). Scanners and drillers: Characterizing expert visual search through volumetric images. *Journal of Vision*, *13*(10):3, 1–13. doi.org/10.1167/13.10.3

Furneaux, S., & Land, M. F. (1999). The effects of skill on the eye-hand span during musical sight-reading. *Proc Biol Sci*, *266*(1436), 2435–2440. doi.org/10.1098/rspb.1999.0943

JASP Team (2020). JASP (Version 0.14.1) [Computer software].

Kinsler, V., & Carpenter, R. H. (1995). Saccadic eye movements while reading music. *Vision Res*, *35*(10), 1447-58. doi.org/10.1016/0042-6989(95)98724-n

Madell, J., & Hébert, S. (2008). Eye movements and music reading: Where do we look next? *Music Perception*, *26*(2), 157-170. doi.org/10.1525/MP.2008.26.2.157

Marandola, F. (2019). Eye-hand synchronisation in xylophone performance: Two case studies with African and Western percussionists. *Journal of Eye Movement Research*, *11*(2):7, 1-24. doi.org/10.16910/jemr.11.2.7

Murray, E. [Edward Murray]. 2019, November 29. *Edmund Rubbra: The Virgin's cradle song (The Sixteen / Harry Christophers)* [Video]. YouTube. https://www.youtube.com/watch?v=NxqpOw_yVUA

Papesh, M. H., Hout, M. C., Guevara Pinto, J. D., Robbins, A., & Lopez, A. (2021). Eye movements reflect expertise development in hybrid search. *Cogn. Research*, *6*(7), 1-20. doi.org/10.1186/s41235-020-00269-8

Penttinen, M., & Huovinen, E. (2011). The early development of sight-reading skills in adulthood: A study of eye movements. *Journal of Research in Music Education*, *59*(2), 196-220. doi.org/10.1177/0022429411405339

Penttinen, M., Huovinen, E., & Ylitalo, A. (2013). Silent music reading: Amateur musicians' visual processing and descriptive skill. *Musicae Scientiae*, *17*(2), 198-216. doi.org/10.1177/1029864912474288

Picking, R. (1997). Reading music from screens vs paper. *Behaviour & Information Technology*, *16*(2), 72-78. doi.org/10.1080/014492997119914

Reingold, E. M., Charness, N., Pomplun, M., & Stampe, D. M. (2001). Visual span in expert chess players: Evidence from eye movements. *Psychological Science*, *12*(1), 48-55.

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), 658-673. doi.org/10.1177/0305735615585398

Rubbra, E. (1926). *The Virgin's Cradle Hymn*. Oxford University Press.

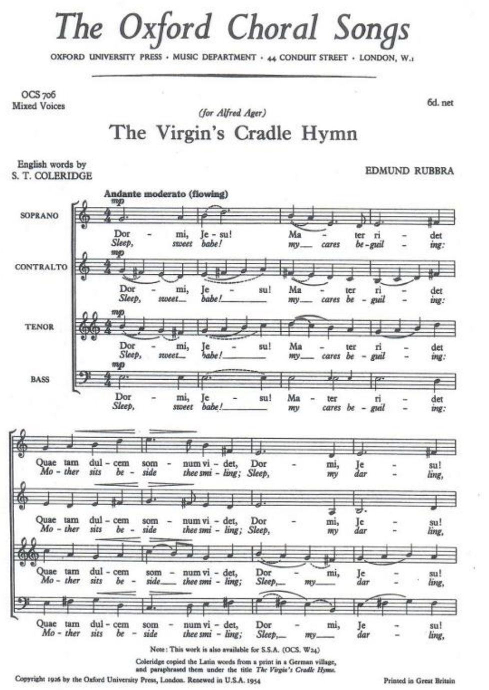
Tobii Pro AB (2014). Tobii Pro Lab (Version 1.145) [Computer software]. Danderyd, Sweden: Tobii Pro AB.

Truitt, F. E., 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.

Wong, Y. K., & Gauthier, I. (2010). Holistic processing of musical notation: Dissociating failures of selective attention in experts and novices. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(4), 541–551. doi.org/10.3758/CABN.10.4.541

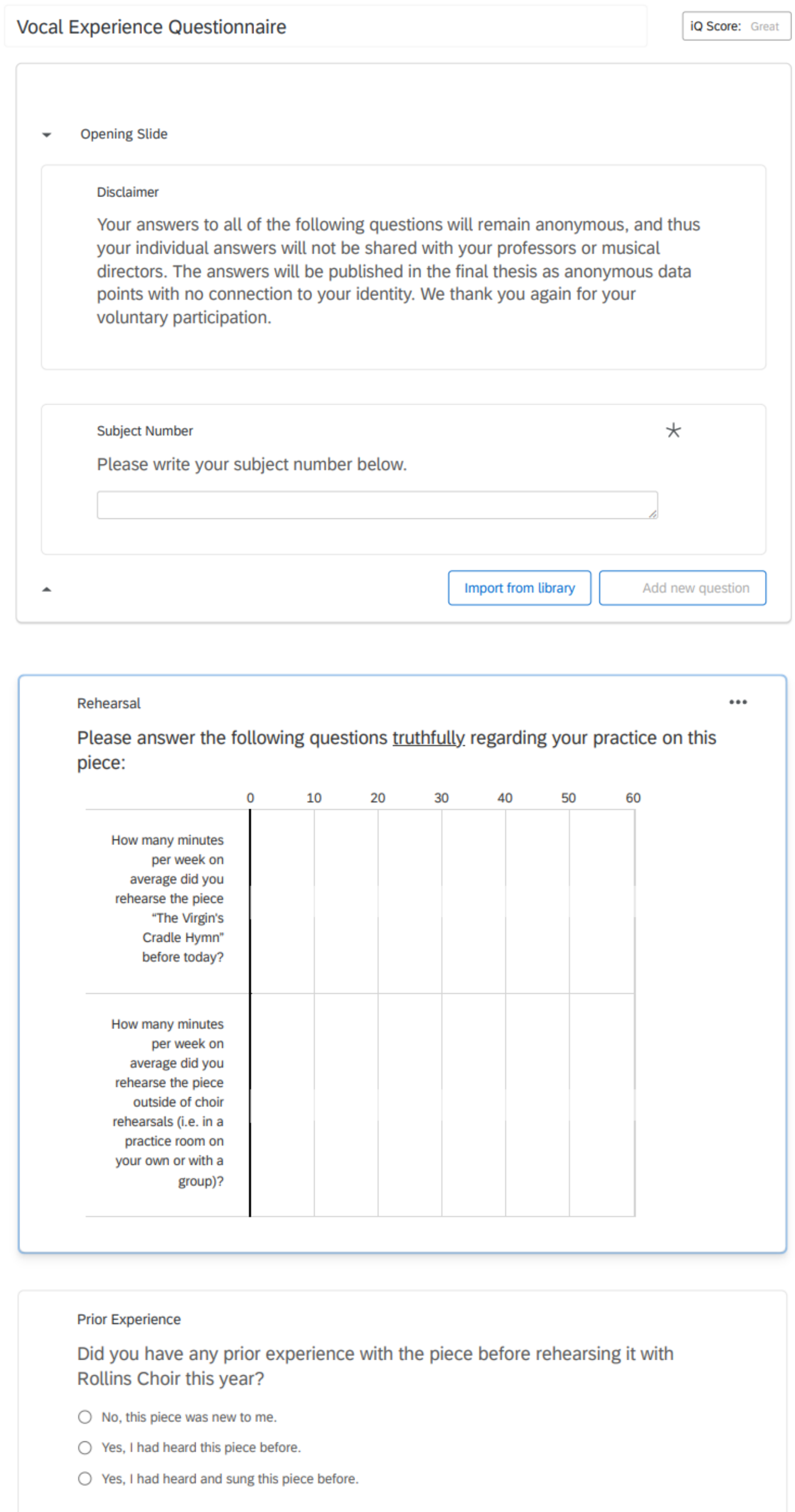
Wurtz, P., Mueri, R. M., & Wiesendanger, M. (2009). Sight-reading of violinists: Eye movements anticipate the musical flow. *Exp Brain Res*, *194*(3), 445-50. doi.org/10.1007/s00221-009-1719-3

Appendix A



Appendix A. These figures are a full copy of *The Virgin's Cradle Hymn* without dividing the different systems into individual slides (Rubbra, 1926).

Appendix B



Vocal Experience Questionnaire

iQ Score: Great

Opening Slide

Disclaimer

Your answers to all of the following questions will remain anonymous, and thus your individual answers will not be shared with your professors or musical directors. The answers will be published in the final thesis as anonymous data points with no connection to your identity. We thank you again for your voluntary participation.

Subject Number *

Please write your subject number below.

Import from library

Add new question

Rehearsal

Please answer the following questions <u>truthfully</u> regarding your practice on this piece:

	0	10	20	30	40	50	60
How many minutes per week on average did you rehearse the piece "The Virgin's Cradle Hymn" before today?							
How many minutes per week on average did you rehearse the piece outside of choir rehearsals (i.e. in a practice room on your own or with a group)?							

Prior Experience

Did you have any prior experience with the piece before rehearsing it with Rollins Choir this year?

- ○ No, this piece was new to me.
- ○ Yes, I had heard this piece before.
- ○ Yes, I had heard and sung this piece before.

General Experience

Please answer the following questions regarding your vocal experience:

	0	4	9	13	18	22	27	31	36	40
How many years have you been singing in choirs, either paid, as class participation, or voluntary?										
How many years have you been singing in Rollins Choir?										
If you are currently taking voice lessons, how many years have you been doing so?										

Demographics

Description

Please answer the following demographic questions:

Class Year

What year student are you (e.g. 4th year/Senior)?

- 1st year
- 2nd year
- 3rd year
- 4th year
- 5th year
- >5th year

Graduation Date

What is your expected graduation date (e.g. Spring 2022)?

	Semester	Year
Graduation Date		

Voice Lesson Level

What level of Applied Voice Lessons are you in?

- I am not taking Applied Voice Lessons
- 101
- 201C
- 201P
- 201P promoting to 301 this semester
- 301

End of Survey

We thank you for your participation in this experiment today!

Please let the researcher know that you have finished.

Appendix B. These are the full list of questions that were included in the Vocal Experience Questionnaire.