Visual Fixation Patterns of Undergraduate Brass and Woodwind Musicians on a Conductor While Playing an Instrument and Reading Music

Douglas James Morin

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Musical Arts

University of Washington

2019

Reading Committee:

Prof. Timothy O. Salzman, Chair

Dr. J. Bradley McDavid

Dr. Steven J. Morrison

Program Authorized to Offer Degree:

School of Music

University of Washington

Abstract

Visual Fixation Patterns of Undergraduate Brass and Woodwind Musicians on a Conductor While Playing an Instrument and Reading Music

Douglas James Morin

Chair of Supervisory Committee:
Professor Timothy O. Salzman
School of Music

The purpose of this study is to explore the visual fixation patterns of undergraduate brass and woodwind musicians on a conductor while playing an instrument and reading music. The factors influencing the ability to view a conductor are explored, including the actions of the conductor and the structure of the written music. A within subjects repeated measures design was employed to explore the questions of participants' visual fixation patterns. Participants (N=22), who were members of auditioned bands at three universities in the Pacific Northwest region of the United States, performed a solo written for this study while viewing videos of an expressive conductor, a non-expressive conductor, and an electronic conductor equivalent. Results showed that participants fixated on the Expressive conductor more than the less expressive conductors. The low overall fixation duration compared to the high percentage of alignment with the conductors' gestures seem to support the use of peripheral vision by participants.

Table of contents

List of Pictures

List of Figures

Acknowledgements

As with all projects of this magnitude, they could not be accomplished without the help and guidance of many wonderful people. First, thank you to Prof. Tim Salzman for all of the direction and discussion concerning the need for this paper, and the feedback about how to improve and present the material. Thank you as well to Dr. Steven Morrison for the advice on constructing the study, navigating the process, and analyzing the data. Dr. Brad McDavid and Dr. Leroy Searle have been invaluable in offering feedback and suggestions for how to proceed.

Several other people have my gratitude for their contributions. Thank you to:

- Scott Rivers and the Human Centered Design and Engineering Department at the University of Washington for allowing the use eye-tracking technology needed to gather data.
- Anita Kumar for help with statistical analysis and the pilot study.
- Lauren Hepburn for help with the pilot study.
- Kaylie Firebaugh for initial design and advice about working with eye-tracking technology.
- Maranda Sianghio for the feedback on the design and writing of this document.
- Dr. Chris Chapman, Mr. Olin Hannum, and Mr. Lewis Norfleet for logistical support.
- The participants who gave of their time to make this study a reality.

Finally, I cannot express the appreciation I have for the support of my family and friends, especially the person to whom this document is dedicated. Not only has the creation of this document taught me about music, it has shown me how much you care.

Dedication

This dissertation is dedicated to my wife, Renee, for her love, constant support, feedback, awesomeness, and encouragement.

Chapter 1: Introduction and statement of the problem

Background

Is the ensemble even watching? Does it matter what I do on the podium? Many frustrated directors mutter similar questions in moments of frustration when the ensemble does not respond to conducting gestures. Previous studies revealed the conductor impacts both the ensemble members and the audience in multiple ways, including shaping listeners' impressions of how the music sounds to guiding the players' interpretation (Luck & Nte, 2008; Thompson, 2012; Sidoti, 1990).

Sight and sound are arguably the most effective methods to communicate a wide array of thoughts with groups of people. While ensemble members coordinate with other musicians through sound (Fredrickson, 1992), conductors I witnessed often chose not to make noise, preferring instead to use physical gestures. Accepting that a goal of music performance is to present sonic art to the audience, it is logical for conductors to rely on sight over sound to prevent interference with the audible atmosphere. In viewing a director's gestures, information transfers from the conductor to both the ensemble and the audience (Morrison, Price, Smedley, & Meals, 2014).

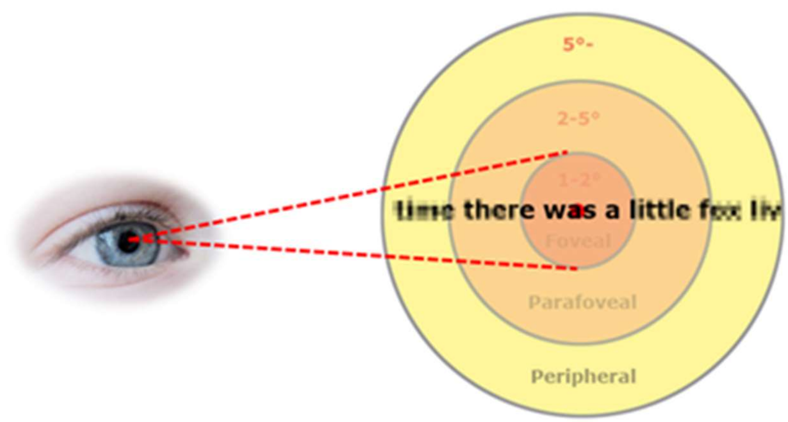
Vision, if the pun may be forgiven, is an often-overlooked sense and interwoven into many facets of musicking. Musicians observe one another to communicate; they watch for cues to start and stop, actions to demonstrate style and tempo, and body language to set moods. Additionally, audiences and performers watch each other for reactions or evoked emotions.

As previously stated, conductors communicate with ensemble members through the sense of sight. Conductors demand ensemble members pay attention to conducting gestures as a standard method to prevent problems, such as inaccurate tempo or inappropriate dynamics. This “watch me” solution might or might not produce the desired correction. The reasons for any lack of improvement could be attributed to either the director or the ensemble members. An astute conductor uses self-evaluation and reflection to assess their conducting gestures and develop motions to symbolize intended musical outcomes. Directors may take multiple approaches toward enhancing their non-verbal skills, including attending conducting workshops, taking lessons with master conductors, receiving input from visiting clinicians, and studying toward advanced degrees. Even if the person on the podium becomes an expert non-verbal communicator, do ensemble members observe a conductor’s gestures? If so, how often do performers view the director and with what part of their visual field?

Visual fields

Researchers typically divide the visual field into three parts: the fovea, the parafovea, and peripheral vision. Each visual field blends from one to another without definitive boundaries. Fovea refers to two related definitions: it is a small depression in the retina of the eye where visual acuity is highest, and it accounts for approximately the center two degrees of the visual field. This area of focus takes in details with clarity and registers small differences, allowing a person to engage in tasks such as reading. The parafovea is the part of the visual field stretching between approximately three to five degrees around the fovea. Objects appearing in the parafovea lack some definition. The primary use of this field of vision is to make small corrections to fixation locations. Peripheral vision encompasses the area wider than the center five degrees of the visual field. While peripheral vision gathers the least amount of detail, it amasses general information, assesses large objects, and registers motion.

Picture 1.1 demonstrates both the approximate degree ranges and how definition is lost in each field of vision. As stated above, the farther away from the fovea an object appears, the less detail a person perceives (Bouma, 1970). The letters appearing in the outer ring, representing peripheral vision, lack the clarity of the letters in the inner circle, signifying the fovea.



Picture 1.1. Visual fields. Source: http://eyetracking.me/?page_id=9

The three portions of the visual field interact to produce a visual pattern. Visual patterns consist of pauses to gather information, known as fixations, and transitions to other spots, called saccades. According to Goolsby (1994), fixations typically last 200 to 250 milliseconds, though these durations may range widely. Saccadic eye movements may be small, such as when a person reads, or large and in combination with movement of the head, as happens when checking traffic at a stop sign. Saccadic suppression, a blurring effect, substantially diminishes the amount of information gathered during a saccade (Rayner, 1998).

Visual patterns and musicking

Many factors affect a person's visual patterns. Of the characteristics that influence the total number and duration of fixations when a person reads music, the largest is the experience of the individual musician. As Gudmondsdottir (2010) discussed, "This is likely due to their (experienced musicians') ability to perceive the musical notation in larger chunks than less proficient readers are capable of. Chunking of information in a musical score depends on the perception of identifiable clusters or entities such as tonal patterns or rhythmic patterns." Multiple studies demonstrate that familiar and predicable patterns, such as chords (Salis, 1980; Waters, Townsend, & Underwood, 1998), musical phrases (Sloboda, 1977), rhythms (Palmer & Krumhansl, 1990; Sloboda, 1983), and tonality (MacKenzie, Vaneerd, Graham, Huron, & Wills, 1986) allow experienced music readers to grasp larger chunks of information. Large leaps, accidentals, and unfamiliar or non-traditional music notation caused slowdowns in reading and changes in visual patterns. (Polanka, 1995).

With all the factors affecting music reading, do instrumentalists have the time or mental capacity to view a conductor and incorporate any of that information? Fredrickson (1992, 1994) published findings concerning the alignment of an individual player sight reading music with a pre-recorded video of an expressive conductor along with an audio recording from a wind ensemble directed by the same conductor. A quarter of the way through the piece, the researcher removed either the audio, the video, neither, or both elements while each participant continued to play to the end of the piece. Fredrickson then evaluated the alignment of each individual participant's performance with the ideal version of the music. Fredrickson made a secondary observation during the study - participants looked directly at the conductor for just under 28% of the total exercise. While this observation seems to address my initial question of "is the ensemble watching," the musical components of the solo used in the study had a direct bearing on the result.

Frederickson selected the music, *Irish Tune From County Derry* by Percy Grainger, due to its "rhythmic and technical simplicity." Appendix A of his 1992 dissertation contains a sample of the music given to the participants. A review of this piece reveals the range was a major eleventh, only two notes were preceded by an accidental, most of the motion was either by step or arpeggio with only six leaps larger than a fourth, and no tempo changes occurred until the very end. An unknown is if the conductor chose the rubato suggested by the "slowish, but not dragged, and wayward in time" tempo indicated on the score. As Chapter 2 will demonstrate, each of these choices directly impacts a musician's visual patterns when reading the music.

Fredrickson addressed sight-reading, an event that by definition only occurs the first time a musician views a piece of music. On the opposite end of the performance spectrum, Byo and Lethco (2001) observed a dress rehearsal, an event which typically occurs at the end of the rehearsal process. They video recorded high school band members from the front, making sure the musicians' eyes remained clearly in the picture. The researchers gathered information about when band members shifted their gaze to the conductor during two pieces, one fast and one slow. Instead of measuring fixation durations, as Fredrickson did, Byo and Lethco categorized the data in “looks per measure (lpm).” For the slower piece, marked Lento, the participants (N = 12) viewed the conductor at a rate of .89 lpm; the faster piece registered a rate of .07 lpm. Due to the different measurement rubric, a direct comparison to Fredrickson’s results is not possible.

Technological advances

Both Fredrickson and Byo and Lethco used standard video cameras facing the participants. In the years since these studies, technology has advanced, allowing computer software and eye-tracking hardware to more precisely track fixation points. The incorporation of infrared cameras facilitates mapping of fixation points by tracking infrared light reflected off the fovea. The reflection is calibrated to known coordinates on a screen. The result of the association between the reflected light, the time it occurs, and the location on the screen the eyes fixate is the visual pattern.

However, the technology has limitations. Science has yet to create a method to directly measure the part of the visual field to which a person gives the most attention. Objects farther from the center of the field of vision become increasingly blurred and less focused. Citing multiple studies, Findlay and Gilchrist (2003) built an argument to support the theory of areas of (visual) attention. The premise of this theory is that attention to any component of the visual field is an overt and purposeful action, meaning information is intentionally gathered through the parafoveal and peripheral fields of vision.

Purpose and need of this study

While research has looked in to the responses to conductors, I cannot find a study that details the visual fixation patterns on the conductor, nor one that addresses the use of visual fields when viewing a conductor. Additionally, most research that measures visual patterns centers on either sight-reading or performance. Yet bands spend lengthy rehearsal time between those two events. My study mimicked a point in the rehearsal process after sight-reading but before an individual reaches a performance ready level.

The purpose of this study was to investigate the question “Do members of an ensemble watch a conductor” by measuring the visual fixation patterns of woodwind and brass players on a conductor while reading music and playing an instrument. As an update to previous research, this study incorporated modern eye-tracking technology to gather detailed data concerning visual patterns.

The data allowed further analyses regarding areas of attention. As an athlete can accurately complete skill drills with a blurred fovea (Ryu, Abernethy, Mann, Poolton, & Enns, 2015), could a musician incorporate the actions of a conductor without shifting the fovea? For instance, if the number of fixations is low, or even zero, yet the person stayed aligned with the tempo of the conductor, it would lend additional support to Findlay and Gilchrist’s (2003) concept of areas of attention. Conversely, an inverse relationship between the number of fixations and number of misaligned measures may suggest that conductors should wait for musicians to become comfortable with the music before conducting the ensemble. Additionally, do a conductor’s expressive gestures, or lack thereof, factor in to the musicians’ visual pattern?

Chapter 2 explores the visual and mental components affecting a musician when reading music. The same chapter also considers the impact of a conductor's actions. Chapter 3 presents the methodology used to design the study. Chapters 4 and 5 report the results of the study and seek to place the results in context, respectively.

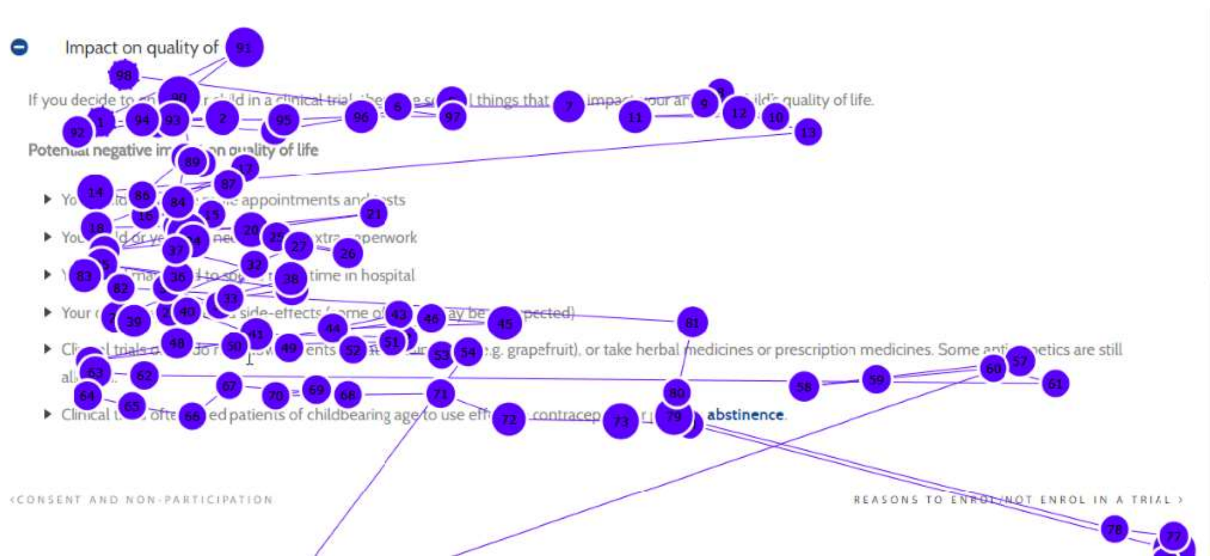
Chapter 2: Background and Literature Review

The purpose of this study is to investigate the question "Do members of an ensemble watch a conductor" by measuring the visual fixation patterns of woodwind and brass players on a conductor while reading music and playing an instrument. Multiple components factor into this question, including how individuals read, the effect of a conductor's actions, and an instrumentalist's ability to respond to a conductor. Since reading music notated in the Western tradition resembles reading printed English words, the exploration of visual patterns begins with studies about comprehending words.

Visual patterns when reading words

When people read words, visual fixation locations principally travel from left to right. According to Goolsby (1994), a typical fixation when reading lasts 200 to 250 milliseconds, though the range may vary widely. The fixations land generally in the middle of words, then jump, or saccade, to the next set of letters. The mean saccade while reading is 7 to 9 letters. When a saccade does not land on the most beneficial spot, a correction is made. Backtracking, a saccade to the left, and adjusting a misplaced fixation when moving from the end of a line to the start of the next line, are the two most common forms of corrections (Findlay and Gilchrist, 2003).

The visual patterns of fixations and saccades are represented by a gaze plot map. Picture 2.1 shows a typical gaze plot map for an individual reading printed word. The filled in circles represent fixations. The larger the circle, the longer the fixation. The number in each circle signifies the order of occurrence. The thin lines connecting the circles denote saccades. In this example, the first fixation is near the word “decide,” which is the first large word of the first full line. The reader then skipped a couple small words, to fixate on the next major word. The end of the first full sentence shows corrections, as fixation 9 is to the left of 8 and fixation 11 is to the left of 10. Based on the placement of fixations 11, 12, and 13, the reader needed to reread the end of the sentence. Note that not all fixations land on a word. Since vision is a field, the reader likely gathered the meaning and progressed through the line without the need to fixate on the center of each word.



Picture 2.1. Gaze plot map when reading words. Source: https://www.researchprotocols.org/2018/5/e119/

The amount of information taken in depends upon the linguistic knowledge, vocabulary, and experience of the reader, not the visual field. Rayner (1986) suggested the perceptual span does not differ from beginning to experienced readers, but rather the difficulty of text in relation to the level of reader causes differences in reading speed. While novice readers take in the same number of letters as an experienced reader, the novice would linger on unknown words and shorten saccades, often with a refixation on the same word. This information suggests an experienced reader created the gaze plot map from Picture 2.1.

The spacing between words influences the size of a saccade. Drieghe, Brysbaert, and Desmet (2005) noted the differences in saccade size and fixation placement produced by extra space between words. Take this sample sentence:

Additionalspacebetweenwordsreducessaccadetimeforreaders.

The lack of space between words increases the amount of time needed to read this sentence. A gaze plot map would likely show fixations in close proximity, several corrections indicating a search for known words, and larger circles representing increased fixation durations. The mental processes affecting the lack of spaces will be discussed later in this chapter. Here is the sample sentence with one space between each word:

Additional space between words reduces saccade time for readers.

The spaces create a guide to a chunk of information. The recognition that each chunk is one word allows the reader to process the combination of letters as one item. This creates shorter fixation durations and larger saccades. On average, a person would read the sample sentence with spaces more quickly than the sample without spaces. An experienced reader knows the spaces define a chunk of information and therefore place a fixation in the middle of words instead of on the first letter. With the lack of space, a reader needs to allow for processing time to determine which letters belong together as they develop the meaning of the combination of letters. Additionally, the spaces seem to be guides to allow for more accurate saccades to the next chunk of information. Drieghe, Brysbaert, and Desmet took the experiment farther and found a significant reduction in reading time for words separated by two spaces.

Visual patterns when reading music.

Studies about reading music produce similar results to those concerning reading words. As seen in Picture 2.2, a gaze plot map taken while reading music shows a similar pattern to reading words. Again, the circles represent fixation points, with the larger the circles representing longer durations. Many of the numbers proceed left to right, with a few regressing from right to left.

The visual patterns between silently reading music and written word are akin; however, music reading is often paired with a further element, an audible performance. To align the two activities, some researchers asked the participants to read out loud. Eye movements differ when reading silently versus aloud. Levy-Schoen (1981) found that mean fixation durations increased when reading out loud. He also noticed that the eyes tended to fixate ahead of the vocalized word.

Picture 2.2. Sample gaze plot when reading music.
Source: https://www.lindsayvickery.com/blog/eye-tracking-ii

Studies concerning musician’s vision often select piano players as participants. The reasons for these choices are typically due to the ease of aligning the key strokes from a MIDI keyboard with the display on a computer screen. These studies reveal insight into the visual and mental processes undertaken by musicians to turn written music into sound.

Goolsby (1994) found that piano players with more experience had shorter fixation durations when reading music and a greater number of saccades. Less experienced players fixated for longer periods of time on the notes. This matches the comparison of experienced versus inexperienced readers, where the novice readers dwelled on unknown words longer. In a study closely aligned with reading words, Waters and Underwood (1998) compared the visual patterns of two distinct levels of music readers without the participants performing the music. The results showed that expert musicians used shorter fixations and larger saccades when sight reading music.

An additional outcome of Goolsby's research demonstrated that experienced music readers did not look directly at each note. As visualized in Picture 2.2, participants focused on a spot near a group of notes. Similarly, when reading words, fixations tend to land in the middle of a word, not at the beginning. Experience in reading notation allows an enhanced ability to recognize patterns and take in larger chunks of information. This concept applies to both pitch and rhythm. Polanka (1995) noted that participants recognized patterns of triads as a unit faster than notes moving in stepwise motion. In analogous findings, Salis (1980) and Waters, Townsend, and Underwood (1998) reported that participants read chords faster than random notes. Sloboda (1977) tested the ability to recognize musical phrases while MacKenzie et al. (1986) studied the role tonality plays in music reading. In both studies participants increased reading speed for both expected musical phrases and tonal passages.

A distinct difference between reading music and written word is tempo. Music has a defined pace; individuals determine their own speed for reading words. Kinsler and Carpenter (1995) tested the effect of tempo on saccades and fixation duration. The results of the study found that as the tempo of the music increased, the time between fixations decreased. This effect applied to all participants, regardless of experience level.

Penttinen, Huovinen, and Ylitalo (2015) measured fixation patterns of adults while playing piano and reading music for known nursery rhymes at a set tempo. More proficient players fixated for shorter periods of time than less experienced players. Then the study introduced unexpected notes which were not in the original tune. The surprising occurrences caused longer fixations with smaller saccades in the process of music reading for all participants, regardless of experience level.

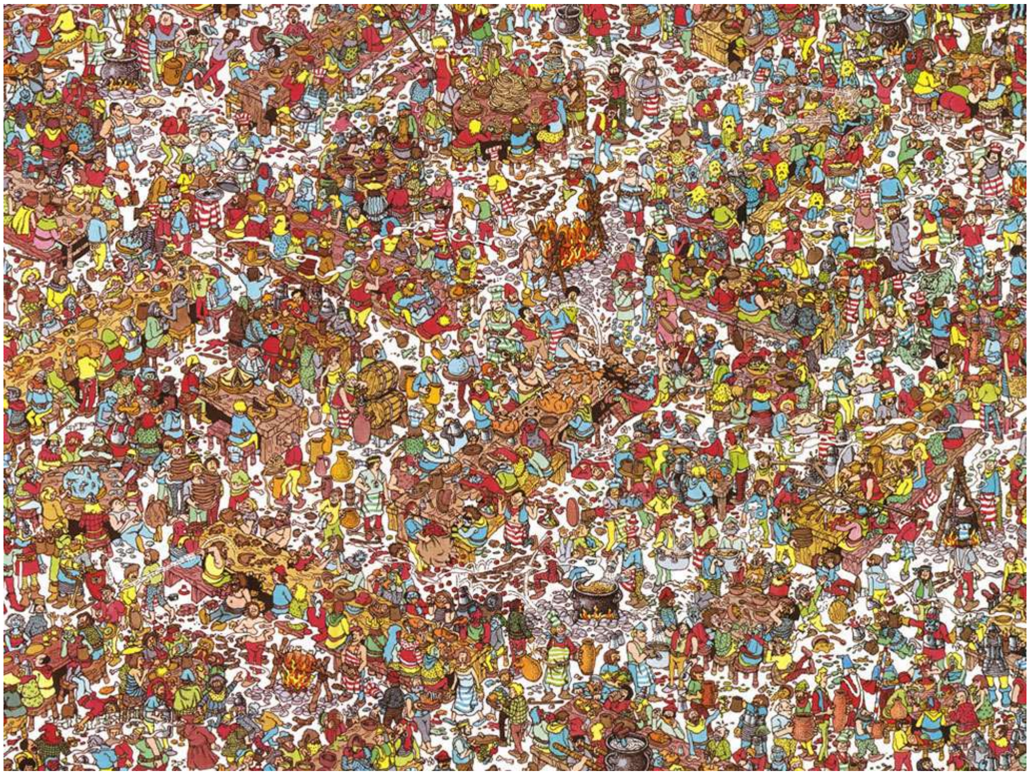
Wurtz, Mueri and Wiesendanger (2009) concluded that the results for violinists comparatively matched previous study results involving piano players. The researchers ascertained that experience is mitigated by the structure of the music; musicians reduced their saccade lengths when reading a more complex piece of music.

Western music relies on both a notation for pitch and symbols representing pitch durations, known as rhythm. Accurate reading of rhythm depends on the reader's ability to mentally construct and reproduce a temporal pattern. Gudmundsdottir (2010) stated "…the decoding of timing information or reading of rhythm is no less important in music reading than decoding of pitch." Studies by Palmer and Krumhansl (1990) and Sloboda (1983) indicated that musicians rely on internal mental representations of musical meter as they perform rhythms. Conversely, unknown rhythms or unexpected notation, such as non-standard beaming, slow down music reading.

Processing information

The amount and relevance of the information in the visual field plays a part in whether the information is incorporated or ignored. The larger the amount of relevant information, the more crowded the visual field becomes. An example of a crowded visual field known to many individuals is a “Where’s Waldo” search puzzle. In these conundrums, the character of Waldo wears a shirt and hat with horizontal red and white stripes along with blue pants. As shown in picture 2.3, this outfit stands out against a background of non-matching colors. However, the Waldo search puzzles include many items and characters approximately the same size as Waldo with both red and white stripes and solid blue, as demonstrated by Picture 2.4. These factors add up to a crowded field of vision. Bouma (1970) first described this concept as the critical spacing needed for recognition of objects. Critical spacing grows in proportion to eccentricity, the distance of the target object from fixation.

Picture 2.3. Waldo alone. Source: http://waldo.wikia.com/wiki/Waldo



Picture 2.4. Waldo in a crowd.
Source: http://www.sienanews.it/cultura/buon-capodanno-si-ma-quale/

Crowding is a function of the brain's ability to process information. If something is deemed irrelevant, the brain quickly discards the information. Schwartz, et. al. (2005) used fMRI technology to test if the information presented in the peripheral vision affected the processing centers of the brain. Schwartz and colleagues displayed images of a flashing checkerboard in the peripheral vision, which had no relationship to the mental task presented in the fovea. In general, the results showed that the more concentration on the central task needed, the less mental response noted in a participant's brain to the extraneous information presented in the peripheral vision.

The ability to include or disregard information is a function of experience. Wong and Gauthier (2012) tested the hypothesis that, "with extensive experience, music-reading experts have acquired visual skills … resulting in higher music-reading fluency.” In a control activity, all participants registered similar response times to identifying mundane objects unrelated to music in a crowded visual field. In the music reading activity, the participants with extensive music reading fluency displayed a smaller crowding effect. Visual spatial resolution improved specifically for objects associated with perceptual expertise.

Sloboda (1985) asked pianists to sight read a piece of music. At a random point, the music disappeared, but the pianists continued performing. Experienced pianists kept accurately playing longer after the music disappeared, especially when the music fit expectations and patterns familiar to the individual. This indicated the experienced participants processed a larger amount of information, implying they perceived it as a chunk of information. Truitt, Clifton, Pollatsek, & Rayner (1997) noted that visual window size did not interact with music reading skill, meaning that similar to reading English letters, the number of notes viewed did not change based on the experience level of the reader.

Taking these other research findings into account, it is not surprising that participants in Fredrickson’s (1992, 1994) study, previously referenced in Chapter 1, could spend so much time looking at the conductor even while sight reading. Frederickson selected the music, *Irish Tune From County Derry* by Percy Grainger, based on “rhythmic and technical simplicity.” The participants were college undergraduate students who proved their capabilities through an audition into an advanced concert ensemble. Few unexpected events occurred, as most notes moved tonally by step or arpeggio with an accidental preceding only two notes. The rhythms incorporated mostly quarter notes, one of the first note lengths learned by most musicians. This meant that each musician stood a better chance of viewing and processing the music quickly. It is likely that the participants high level of experience allowed them to chunk patterns of notes, allow more time and processing ability to focus on the conductor.

Integrating conductors' actions

With the numerous factors influencing the processing of objects in the visual field, how much information could instrumentalists incorporate from a conductor while playing and reading music? Cofer (1998) analyzed seventh grade band students' performances on instruments when responding to a conductor. All students received practice time with the music before participating in the data gathering phase of the study. The treatment group of students received additional instruction on the meaning of conducting emblems outside of regular ensemble rehearsals. The control group received no additional instruction. Students in the treatment group identified conducting gestures at a significantly higher rate both while playing instruments and during a pencil-and-paper test.

Thompson (2012) extended Cofer's research to high school aged string players. Thompson also divided the students into two groups. The treatment group received five instructional sessions focused on conductor gestures outside of the normal class time. While the students in both groups responded more successfully to a "pencil and paper" post-test, the participants in the treatment group correctly responded to a larger number of expressive conductor gestures than did participants in the control group.

Sidoti (1990) measured students' playing for expressive sounds. Participants performed to videos of a conductor who either displayed or did not display gestures reflecting the expressive markings on the music. Students accurately played a higher percentage of musical marks when the conductor's gestures matched the musical markings.

The experience level of a musician is the key factor in predicting how well they will synchronize their actions with a conductor's gestures. Luck and Nte (2008) asked participants to tap in time with simple conducting gestures while the researchers applied treatments in order to collect data about synchronization accuracy. These factors were "(1) the radius of curvature with which the beat was defined; (2) the experience level of the conductor; and (3) the experience level of participants." Though all participants responded to the gestures in some way, results indicated that only participants' previous experience affected their synchronization ability.

Characteristically, musicians use both their sense of hearing and sight while playing an instrument. Clayton (1986) sought to find the relative contribution of the factors of musical timing, which he defined as the ensemble, the conductor, the score, and an individual's own internal sense of timing. During the study, Clayton placed instrumentalists in various scenarios, including adding or removing the sound from control recordings and changing participant's ability to see a conductor. Among the conclusions of this unpublished dissertation, Clayton reported that having a conductor in view reduced the "window of the initiation" of notes, meaning participants were more likely to synchronize the start of notes with the recording.

Could a musician use peripheral vision to view a conductor? Though objects in peripheral vision are not as clear as in the fovea, people are capable of determining the identity of some items and registering motion. Calvo and Beltrán (2013) found participants could categorize faces with happy expressions faster than angry, fearful, sad, and neutral faces.

Traschütz, Zinke, and Wegener (2012) noted deceleration detection becomes superior to acceleration detection in peripheral vision. Foveally, average detection thresholds are lower for accelerations than for decelerations. Using basketball players, Ryu, Abernethy, Mann, Poolton, & Enns (2015) introduced increasingly greater fields of blur into the vision of participants. Those with higher skill and experience were able to complete tasks despite a blurring of peripheral vision.

WÖllner (2008) created a study to determine what part of a conductor's body displayed the most expressive information. The sample videos included an unaltered view of a conductor, only the arms, the face alone, and a blurred view of the conductor. Participants rated the two videos that displayed the conductor's arms clearly as containing more information than the two videos that removed or blurred the view of the arms. WÖllner did not employ any form of eye tracking technology to address the question of where the participants fixated on the unaltered conductor video.

Purpose of this study

This brings the conversation back to the original question, "Is the ensemble watching" the conductor? While previous research suggests musicians watch, a variety of factors affect the ability to do so.

The purpose of this study is to investigate instrumentalists' fixations on a conductor while playing an instrument and reading music. This study seeks to replicate the rehearsal time between sight-reading and performance, a period of time yet to receive attention. Several other questions stemming from the initial query will also be considered. Does the expressiveness of a director factor into the visual fixation patterns on a conductor by an instrumentalist when a performer is presented with challenging music? Could a conductor be replaced by an electronic facsimile? Is the fixation count or fixation duration proportional to the alignment of the instrumentalist with the conductor? Does a participant need to foveally view a conductor to stay aligned with the beginning of each measure? What part of a conductor's body do participants chose to focus?

Chapter 3: Methods

The purpose of this study was to investigate the visual fixation patterns of woodwind and brass players on a conductor while reading music and performing on their instrument. This chapter describes the characteristics of the participants, materials, procedures, and analyzation process for the study.

Study design

To address the questions posed in the creation of this study, I employed a within subjects repeated measures design. The experimental interface included two videos of a human conductor displaying different levels of expressivity, one video of an electronic facsimile displaying a conducting pattern without a human visible, and one control slide containing only printed word instruction. Participants played their chosen instruments while reading the given solo while watching the video interface. The presentation order was randomized for all participants. Eye movements were tracked using eye-tracking technology, then analyzed for fixation counts and fixation duration.

Participants

I incorporated several steps to narrow down which musicians would participate in the study. Since I was researching the visual patterns of an individual on a conductor, participants needed to partake in conducted ensembles. To control for differences between the three types of ensembles - bands, choirs, and orchestras - that typically feature a conductor, I selected concert band members due to a higher available population count. I discounted percussionists based on a tendency I have observed to look directly at pitched instruments while playing, which would affect the visual pattern. Additionally, percussionists stand to play, which may have unknown effects on their visual patterns. This left woodwind and brass players as the potential participants.

I recruited participants (N = 22) by contacting the Director of Bands at three universities in the Pacific Northwest region of the United States. Upon granting permission, each director received a recruitment email to forward to the eligible ensemble members. The email introduced the primary researcher, identified who gave permission to send the email to the students, explained the purpose of the study, detailed the criteria to become a participant, and gave an overview of what each participant would do if they elected to participate. All participants met the following eligibility criteria: at least 18 years of age, undergraduate student, accepted by audition into a concert ensemble, and play a woodwind or brass instrument. The email clarified that participation was voluntary and that the decision to take part had no connection to any class at any university. Potential participants filled out a short web-based form to indicate interest and available times. The primary researcher confirmed a specific time with each participant and answered all initial questions through email communication.

Eye-tracking camera and software

The best method to track visual fixation patterns is infrared eye-tracking technology. A series of calibrated cameras report where a person's fovea aims based on reflected infrared light. The decision to use this technology stemmed from the software's ability to incorporate a projection of the conductor videos. The Tobii Corporation created both the eye-tracking software and hardware. The hardware, the Tobii Pro X3-120, features three built-in infrared cameras with a collection rate of 120 Hz. The best software to utilize the capabilities of the Tobii Pro X3-120 is the Tobii Studio Pro, version 3.4.8. Niehorster, Cornelissen, Holmqvist, Hooge, and Hessels (2017) rated this system as highly effective in ability to regain eye movements after a participant looked away from the target area. The effective distance for participants to sit from this eye-tracking system is between 45 and 85 centimeters. The optimal distance window stretches from 55 to 75 centimeters from the eye tracker to the participant. Though originally intended for use with computer screens, pilot tests proved the Tobii eye-tracking system could reliably gather data when a projector displayed images onto a wall mounted screen.

Tests during the pilot study initially began with two eye tracking systems operating concurrently. One system recorded fixation data on the conductor videos; the second system logged fixation data as participants read the music. However, the simultaneous use of the two eye-trackers produced false data. Fixation points appeared to jump randomly around the screen. Apparently, infrared light from one set of cameras bounced off the fovea and both units recorded the reflection. Since the purpose of the study aimed to address when participants focused on the conductor, the eye-tracker recording fixations on the music was removed.

To account for differences in participant height and chosen stand placement, the cameras resided on an adjustable stand, allowing the camera to reside just over the top of the music.

Novel solo

I wrote a novel solo, available in Appendix C, specifically for this study to approximate the technical challenges typically found in an auditioned college ensemble. This work totaled 22 measures at a tempo of 108 beats per minute. The piece was scored in concert A-flat. Finale Notation Software transposed the solo into the appropriate clef and key signatures for each potential participating instrument. The primary researcher then displaced some pitches by one octave to place phrases into appropriate ranges. The required range for each instrument stayed within the interval of a twelfth. Rhythms included half, quarter, dotted quarter, eighth, dotted eighth, and sixteenth notes. The beams fit the traditional manner of appearance. A tempo mark appeared at the beginning of the first line; no other musical marking existed on the page.

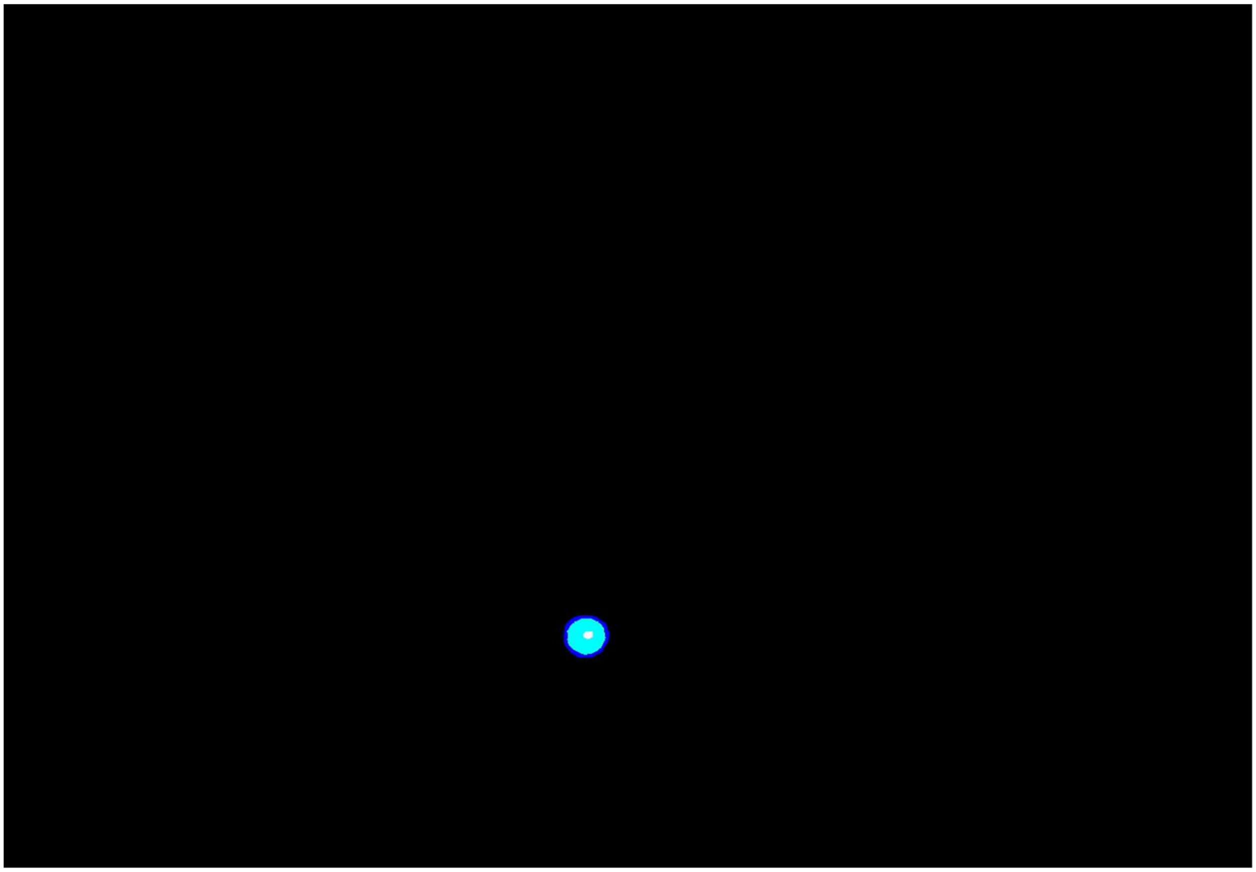
The composition contained elements identified by previous research as likely causes of fluctuations in instrumentalist's fixation patterns. Syncopation happened in measures 4, 8, 10, 17, 18, and 20. Measures 6, 16, 17, and 19 involved accidentals. Pitches modulated from recognizable patterns, including scales and arpeggios, to large leaps and unexpected motifs.

Conductor videos

The study contained three videos of conductors. When creating each video, the camera sat in the middle of the second row of a simulated concert band set up. While taping each video, a midi recording of the solo played to ensure identical tempos. Video editing software removed all sound during post-production and converted all videos to a file format compatible with Tobii Studio Pro.

Two videos featured a human conductor. In one (referred to hereafter as the "Non-expressive conductor video") the conductor showed a static and unresponsive facial expression while maintaining eye contact with the camera for the entirety of the video. The baton and left-hand mirrored each other in an unchanging focal point 4-4 conducting pattern. The other video of a human conductor (henceforth referred to as the "Expressive conductor video") displayed a conductor who changed facial expressions while maintaining eye contact with the camera. The baton employed a focal point 4-4 conducting pattern which changed size to indicate different dynamics. The left hand operated independently, portraying a wide array of conducting gestures throughout the video. Gestures incorporated in the Expressive conductor video included crescendos, decrescendos, accents, legato patterns, and staccato patterns.

The third conducting video hid the human conductor. Instead, it presented a blue circle of light on a black background, seen in Picture 3.1, (here within referred to as the "Dot conducting video") moving in a focal point 4-4 pattern devoid of style or size changes. To create this video, the same human conductor from the Non-expressive and Expressive conductor videos held a blue lens over a flashlight in their right hand in a dark room. Editing software adjusted the video's contrast to ensure the background turned black so to not reveal the human holding the flashlight. The Dot conducting video contained no visual information outside of the moving blue light.



Picture 3.1. Participant's view of the Dot conductor video

To account for random fixations, participants played the solo to a static instructional screen (herein referred to as the "No conductor instructional screen"). The words displayed on this screen stated, "Please play the solo without a conductor."

For reference purposes through Chapters Four and Five, the Non-expressive and Expressive conductor videos will be referred to as the "human conductor videos" and the Dot, Non-expressive, and Expressive conductor videos together will be referred to as the "conductor videos."

Room set-up

All equipment was set up before the participant entered the room (See Picture 3.2). The primary researcher sat behind a table to the left of the participant's chair. The table contained two laptops (one controlling the eye-tracker and one for general use), the camera to record the conductor videos while participants played, and space for paperwork. The area for the participant included one standard height chair without arms, a stand with the music already on it, and a second stand turned with the lip up on which the eye-tracking camera hardware rested. The eye-tracking cameras were secured to the stand lip with poster putty to prevent inadvertent movement and allow for minute changes in angle. A projector was positioned to the right and forward of the participant in two of the rooms. The third room contained a built-in projector. All projectors displayed the image of the human conductor at the approximate size that a live human conductor would appear. The projector settings and lights in the room were manipulated to allow an optimal and clear view of the conductor videos.

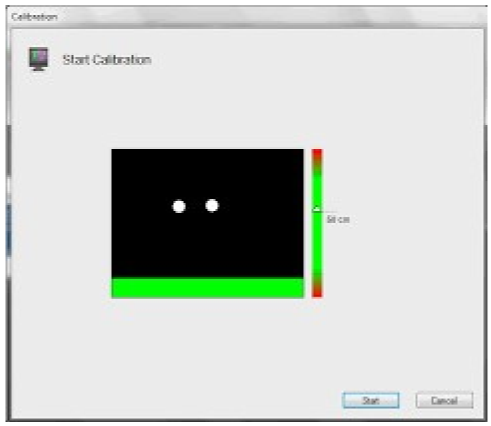
Picture 3.2. Room set-up for the study

Procedures and Data Collection

At the beginning of the appointment the primary researcher greeted each participant, presented two copies of the consent form, see Appendix A, gave each person time to read over the form, and informed each person that they may ask questions at any time. Upon signing the consent form, the participant assembled their instrument and played any desired warm-up with the solo in view on a stand. The researcher invited participants to play through the solo as many times as the individual desired with the request to not memorize the music. Participants who only practiced the solo vastly under tempo received instruction about the actual tempo; the researcher provided no other instructions concerning the music on the stand.

When each participant deemed themselves ready, they received a demonstration of the eye-tracking equipment and procedures for the study. A sample video of a human conductor played on the screen while the researcher explained the conductor's gestures signifying the times when the participant should start and stop performing the music. The essential instruction concerned the conductor indicating four beats in tempo before the start of measure one. At this point the participant adjusted the music stand to the individual's chosen height to view the conductor videos. With the stand height set, the researcher altered the height of the stand with the eye-tracking cameras. A second demonstration video displayed a sample of the Dot conductor video with emphasis on the measure setting the tempo before the participant began playing. The researcher informed the participants that they would play the solo a total of four times, with the conductor videos and the No conductor instructional screen presented in a random order. Participants confirmed they understood the instructions and asked clarification questions.

Next, the eye-tracking cameras turned on and each participant acknowledged they could see the three red lights on the eye-tracking hardware. The software presented a real time graphic of the infrared light reflected by the fovea and captured by the cameras. As Picture 3.3 displays, the fovea are represented by the two white circles on the black background. The green bar under the black field represents the signal strength of the reflected infrared light; yellow meant a weak signal and red represented no signal. The bar to the right of the black field measured the distance between the cameras and the fovea, with the green center representing the most effective distance. The researcher adjusted the eye-tracking hardware as necessary to enhance the effectiveness of the eye-tracking system.



Picture 3.3. Calibration graphic for Tobii Studio Pro

The calibration process began with instructions to put the instrument in playing position, follow the moving red circle on the screen with minimal head movement, and the initial location of where the calibration dot would first appear on the screen. Participants wearing glasses received a request to clean the glasses to improve the response of the system. The calibration proceeded when participants acknowledged their readiness. Some participants required multiple attempts to successfully calibrate the software.

With calibration complete, the researcher manually indicated to the software the participant's readiness and the recording of data began. Title screens before each of the four treatments stated, "Please follow the conductor." The software randomized the presentation order of the conductor videos and the No conductor slide. Following each treatment, the program automatically changed to a screen displaying the message "Thank you." Each participant acknowledged readiness before proceeding to the next video. After the completion of the study, the researcher thanked participants for their contribution and provided a final opportunity to ask questions in person. All participants finished the study in under 30 minutes.

Evaluation and Analysis

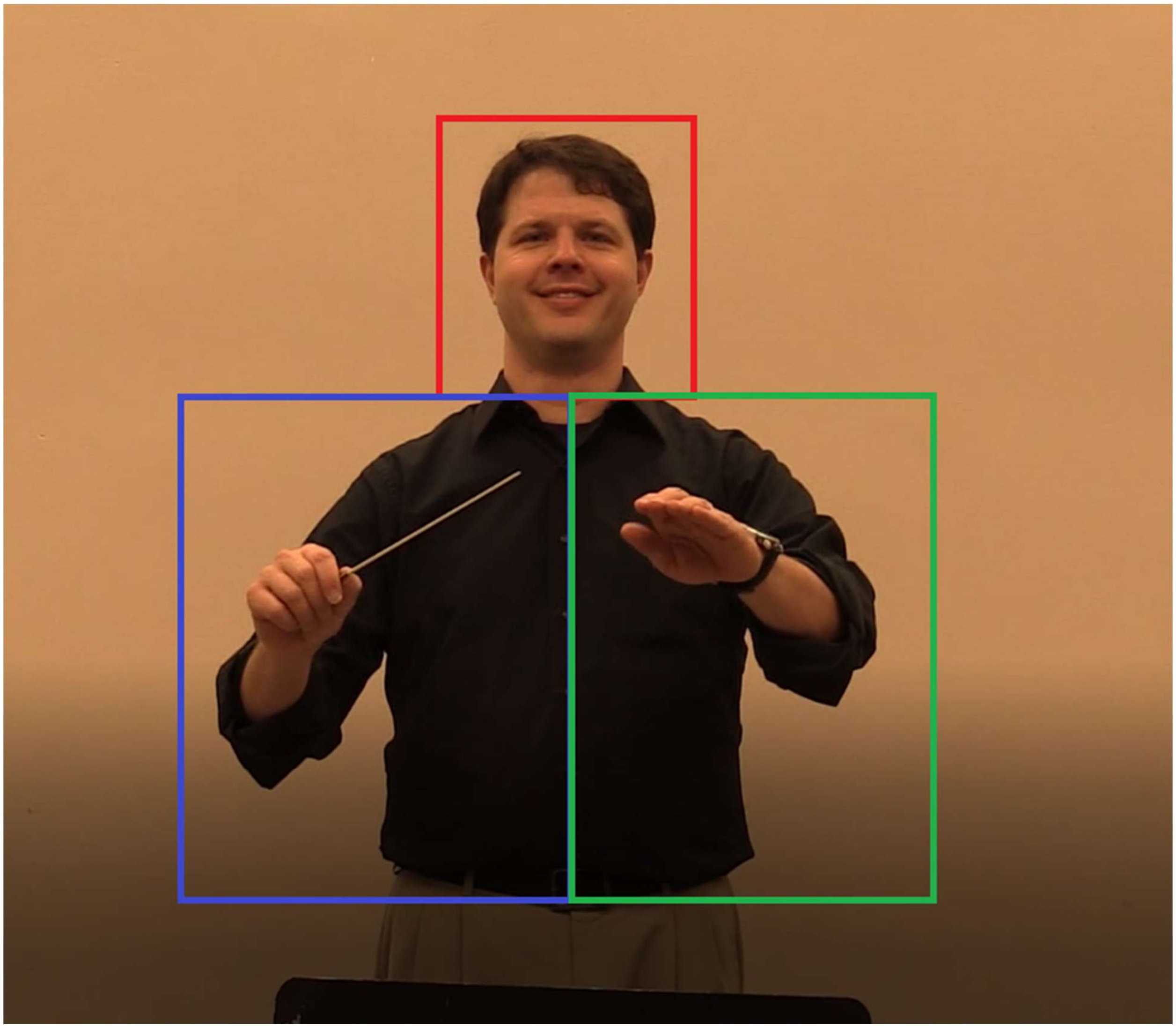
The Tobii Studio Pro software can sort and present data in a variety of ways. The key data points used to address questions concerning the participants' eye contact on the conductor videos were visual fixation counts and durations. To ascertain if all fixations should be equally weighted, the researcher examined fixation locations, the data generated by the No Conductor screen, and the chance for presentation order bias.

The analysis of each conducting video began with totaling the fixation counts and durations from the initiation of the count-off gesture to the cessation of the conductor's movement, a total of 47.5 seconds and henceforth referred to as "full time." Another analysis apportioned the conductor videos into three segments: the three seconds from the initiation of the conductor's movements through the first beat of measure one, the two and a half seconds of the last measure including the cutoff gesture, and the time between the first three seconds of movement and the final measure, which totaled 42 seconds, hereafter referred to as "the middle."

Total fixation counts from each segment revealed the percentage of fixations that occurred in each segment. I used a repeated measures ANOVA to test for differences between the fixation counts of each conductor video. Where appropriate, pairwise comparisons determined differences between the fixation counts of specific videos for each defined time segment.

Calculations of fixation durations divulged the total and mean times for entire videos, as well as during each segment of the videos. I used a repeated measures ANOVA to test for differences between the fixation durations of defined video segment. Where appropriate, pairwise comparisons determined differences between the fixation durations of specific time delineations.

To determine if one location of the conductor's body received more fixations over another, the researcher defined five Areas of Interest (AOI), pictured in Picture 3.4, on the two human conductor videos. The areas of interest include the entirety of the whole screen, the conductor's face, the entire torso, the right side of the torso, and the left side of the torso. All torso areas include space outside the body the hands and baton may travel. The baton tip occasionally traveled from the right box to the left box. The lack of face or body displayed by the Dot conductor video rendered separate areas of interest unnecessary. I used a repeated measures ANOVA to test for differences between the fixation counts and durations from each areas of interest.



Picture 3.4. Areas of Interest on the human conductor videos

Audio and video recordings allowed for tracking of musical alignment. The key moment of alignment for this study was the comparison between when the participant played the first note of each measure and the conductor's indication of beat one. A participant needed to play the first note on the beat one of the corresponding measure for the conductor to be considered aligned. Misalignment occurred when a participant played the first note at any other time than the conductor's beat one. A Pearson product-moment correlation coefficient was calculated to determine if any relationship exists between the number of misalignments and the fixation counts. A separate Pearson product-moment correlation coefficient was calculated to determine if any relationship exists between the number of misalignments and the total fixation durations by each participant.

Several participants skipped notes or extended rhythm durations in a measure to begin the succeeding measure with the conductor's gesture. This action was defined as a correction. A Pearson product-moment correlation coefficient was calculated to determine if any relationship exists between the number of corrections and the fixation counts. A separate Pearson product-moment correlation coefficient was calculated to determine if any relationship exists between the number of corrections and the total fixation durations by each participant.

Chapter 4: Results

The purpose of this study was to investigate the question "Do members of an ensemble watch the conductor?" To answer this inquiry, eye tracking technology collected data about instrumentalists' fixations while playing an instrument and reading music with a conductor. Several other questions stemmed from the initial query. Does the expressiveness of a director factor into the visual fixation patterns on a conductor by an instrumentalist when a performer is presented with music written with numerous complex components? Could a conductor be replaced by an electronic facsimile? Does a participant need to foveally view a conductor to align the first note in each measure with the conductor's gesture and do the fixation counts and durations increase with more accurate alignment?

Basic information

A total of 22 people from three public universities in the Pacific Northwest region of the United States met the following conditions and volunteered to participate in the study. The criteria for participation stated that a participant must be an undergraduate student, over the age of 18, play a woodwind or brass instrument, and be accepted into an auditioned indoor concert ensemble at the respective university. All participants (N = 22) completed the study, including paperwork, in under 30 minutes. The participants self-selected the following instruments to play during their study participation: two chose flutes, five clarinets, three bassoons, one alto saxophone, two tenor saxophones, three trumpets, three French horns, one trombone, one euphonium, and one tuba. Due to the small number, the analysis of data considered all instrument populations together. Individual instruments or instrument families lacked the statistical power to make reliable calculations.

Error detection

To remove concerns of order bias or false data, the primary researcher examined fixation locations, the possibility of inadvertent fixations, and presentation order. The No Conductor instructional slide tackled the concern of false data created by unintentional or random fixations. Participants totaled 216 fixations on the No Conductor instructional slide. Of these, 214 fixations happened in the initial nine seconds - the time participants required to read the words on the slide, raise the instrument, and begin to play. Two separate participants recorded the final two fixations, which occurred at the 1:03 and the 1:26 minute marks. A cross reference with the audio recordings revealed these fixations fell after the respective participants finished playing the excerpt. None of the participants fixated on the No Conductor instructional screen while playing their instrument, which indicated a minimal chance of random, unintentional random fixations throughout the other three conductor videos. The fixation data from the No Conductor instructional slide received no further analysis.

The initial method considered during the pilot study to investigate fixation data called for the creation of a single area of interest (AOI) covering only the area within which the conductor moved. The fixation plots for all three conductor videos revealed fixations outside of conductor's window of motion. Since no participant fixated on the No Conductor instructional screen, it was safe to assume that participants intended all recorded fixations on the conductor videos, even if the saccade resulted in a fixation on an unintended point. Therefore, the AOI expanded to include the full screen. Combining the fixation counts for the Dot, Non-expressive, and Expressive conducting videos, the fixations on the full screen totaled 577.

To control for presentation order bias, I employed a Tobii Studio Pro software setting to randomize the appearance sequence of the three conductor videos and the No Conductor instructional slide. Figure 4.1 depicts the mean fixations per person for the three conductor videos divided by presentation order, not by the specific video. This analysis did not include data from the No Conductor instructional slide. An eye test of the means suggested no effect from the order of presentation. I used a repeated measures ANOVA to test for differences between these fixation counts. Mauchly's test of sphericity was not significant. Results showed no significant main effect between conductor videos, F (3, 45) = 0.027, p = .99, *partial* $\omega 2$ = .002. Based upon these results, the data for each video counted equally in all further analysis.

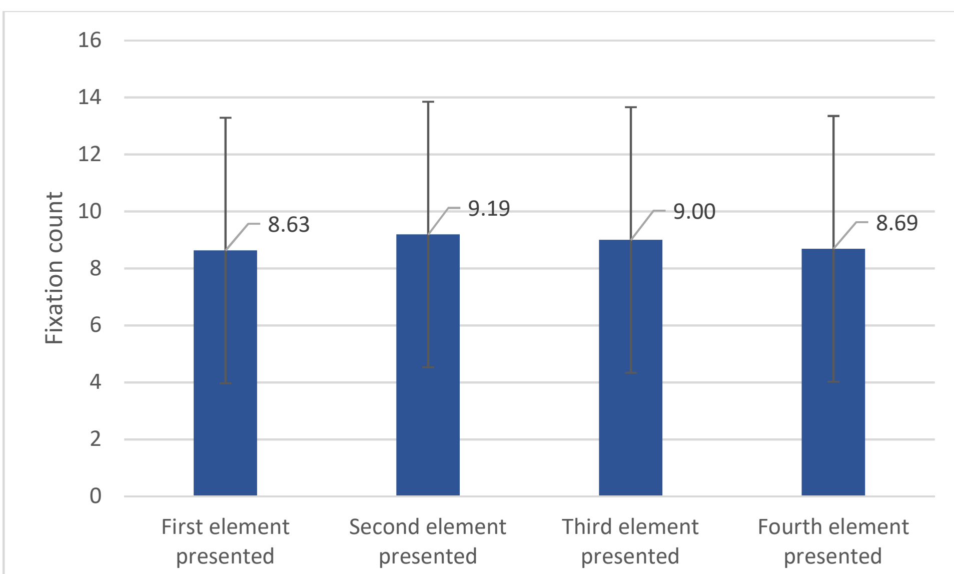


Figure 4.1. Mean fixations based on presentation order

Analysis of data over full time of videos

Seeing no reason to discount any data, analyzation of the fixation counts began with the full time of each conductor video, from the initiation of the count-off to the cessation of the cut-off, a total of 47.5 seconds. As displayed on Figure 4.2, participants (N=22) totaled 148 fixations upon the Dot conductor video, 172 during the Non-expressive conductor video, and 257 through the Expressive conductor video. I used a repeated measures ANOVA to test for differences between these fixation counts. Mauchly's test of sphericity was significant, so the report shifted to Greenhouse-Geisser adjusted *F*-tests to control Type I errors. Results showed a significant main effect on conductor video, Greenhouse-Geisser Adjusted F (1.43, 29.99) = 5.65, p = .02, *partial* ω2 = .17. A pairwise comparison revealed a significant difference occurred only between the Non-Expressive and Expressive conductor videos, p = .03, with the Expressive conductor video totaling significantly more fixations. No other significant differences occurred between the other videos. One person did not fixate on the video for the entire duration of the Dot conductor video. A different person totaled zero fixations on the Non-expressive conductor video.

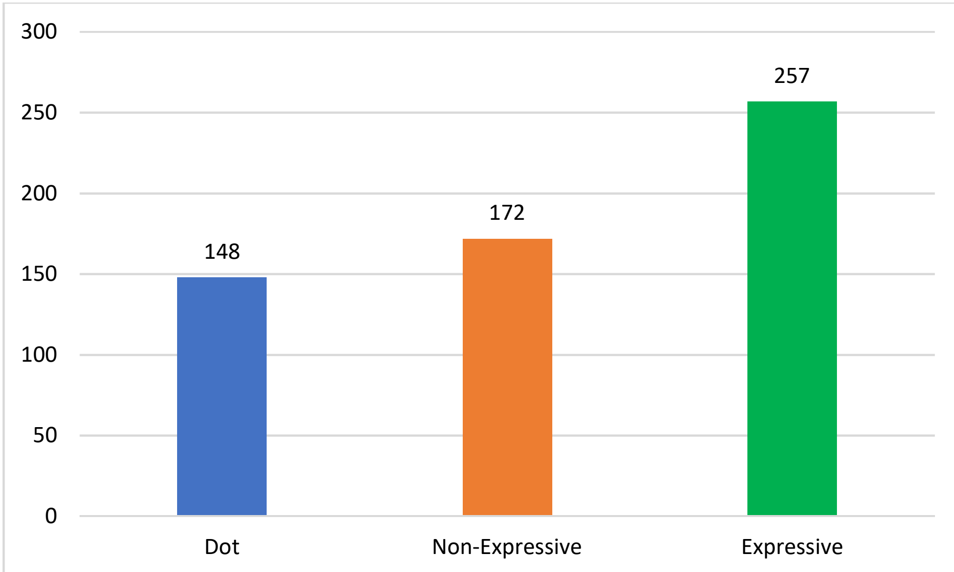


Figure 4.2. Total fixation counts

Fixation durations had a wide range of lengths, best represented by the standard deviation for each mean. Figure 4.3 depicts the mean fixation durations. During the full length of each video, participants viewed the Dot conductor video for a total of 38.96 seconds with a mean of 1.77 seconds (3.73%, SD = 1.24 seconds), the Non-expressive conductor video for a total of 53.12 seconds with a mean of 2.41 seconds (5.07%, SD = 1.63 seconds), and the Expressive conductor video for a total of 66.73 seconds with a mean of 3.03 seconds (6.38%, SD = 2.35 seconds).

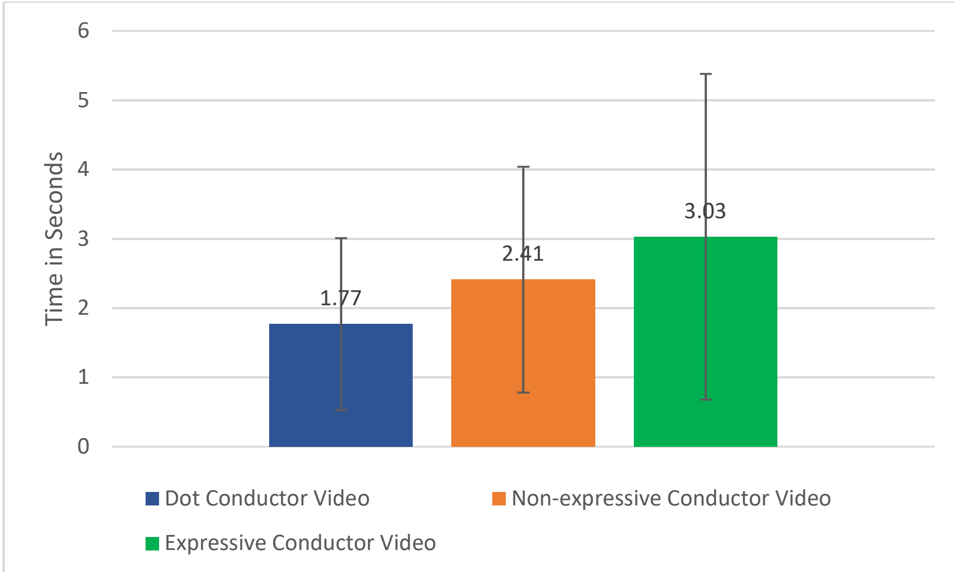


Figure 4.3. Mean fixation durations

I used a repeated measures ANOVA to test for differences between these fixation durations of the full videos. Mauchly's test of sphericity was not significant. Results showed a significant main effect of conductor video, $F(2, 42) = 7.69$, $p = .001$, *partial* $\omega 2 = .23$. A pairwise comparison revealed a significantly longer fixation duration for the Expressive conductor video than the Dot conductor video, $p = .006$. No other significant differences occurred between the other conductor videos.

Figure 4.4 displays the individual participant with the longest fixation length. The maximum fixation duration by an individual participant totaled 4.45 seconds (9.37%) of the Dot conductor video, 6.55 seconds (13.79%) of the Non-expressive conductor video, and 8.00 seconds (16.84%) of the Expressive conductor video.

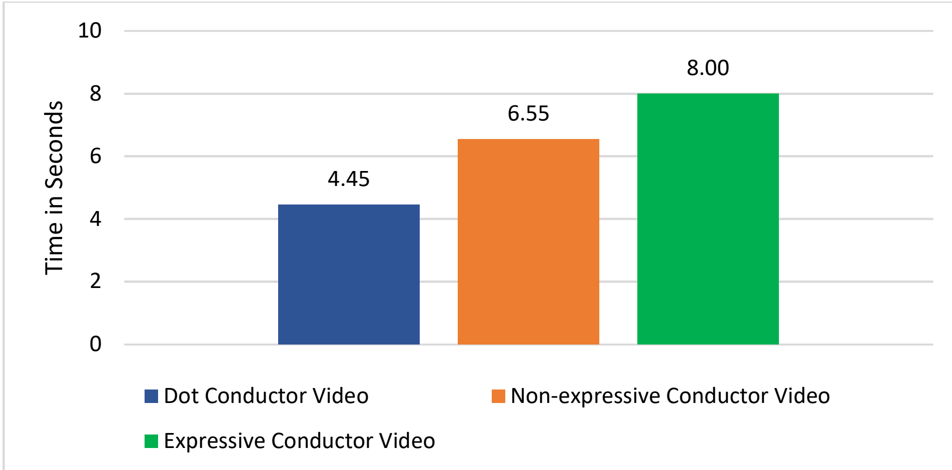


Figure 4.4. Maximum fixation duration by an individual participant

Fixation counts for video sections

Visual inspection of a sliding gaze plot chart showing fixations appear and disappear in real time hinted at an unequal distribution of fixations through the course of each video. Therefore, the data were divided into three sections: the count-off and first beat, totaling 3 seconds (6.32% of the video), the start of the final measure to the cessation of the cut-off gesture, totaling 2.5 seconds (5.26% of the video), and the time in the middle between the first beat and the final measure, totaling 42 seconds (88.42% of the video).

Figure 4.5 demonstrates how the fixation counts when combining the data for the three conductor videos based on time segments. The count-off and first beat received 268 fixations. That equated to 46.45% of the total fixations, a percentage disproportionally larger than the time duration of 6.32%. The proportion of fixations for last measure, 110 fixations (19.06% of the total), was also larger than the ratio of time. The middle section should have equated 88.42% of the fixations. Yet, it only received 34.49% of the total (199 fixations).

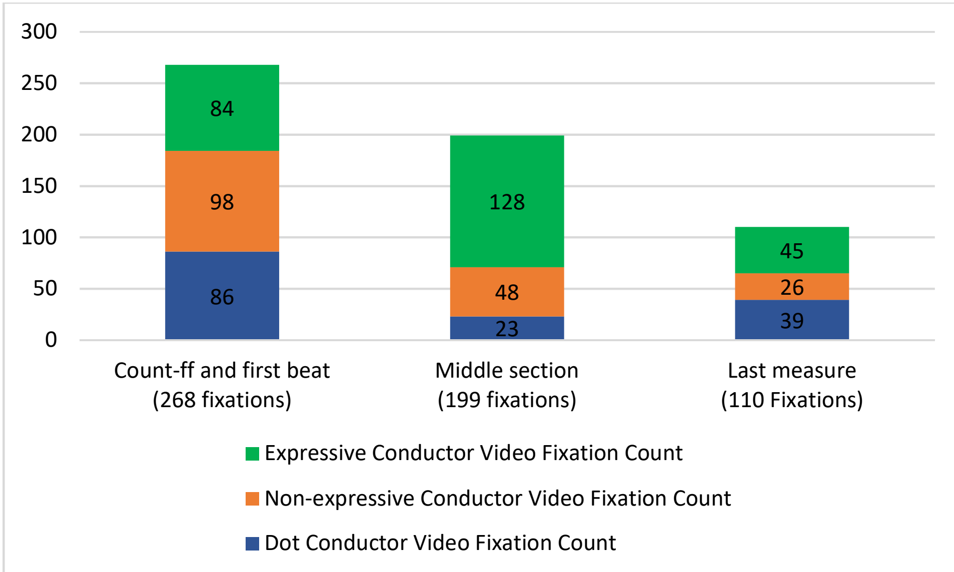


Figure 4.5. Fixation counts by section, all conductor videos

Participants fixated on the Dot conductor video more during the count-off and first beat than at any other time segment. Of the 148 total fixations from all participants, the count-off and first beat registered 86 fixations equaling 58.11% of the total fixations, the middle section garnered 23 fixations (15.54% of the total), and the final measure 39 fixations (26.35% of the total). When comparing the percentages of fixations occurring in each section, visualized in Figure 4.6, the middle section, shown in orange, shrank substantially, while the first section and last measure both grew.

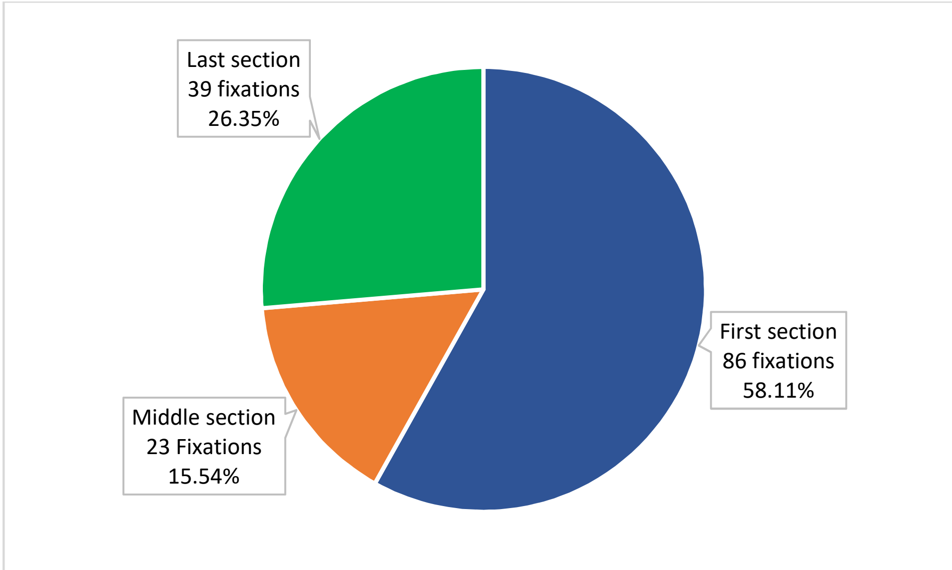


Figure 4.6. Dot conductor video fixation counts and percentages

Breaking down the data for the Non-expressive conductor video revealed different proportions. Participants documented a declining number of fixations from beginning to end. Of the 172 total fixations, the count-off and first beat registered 98 fixations (56.98% of the total fixations); the middle section garnered 48 fixations (27.91% of the total fixations); the final measure recorded 26 fixations (15.12% of the total fixations). These percentages are different than the proportional breakdown of the time. The beginning, blue, is well over half of Figure 4.7. The middle section, displayed in orange, encompasses a larger portion of the pie for this video than the corresponding section from the Dot conductor video.

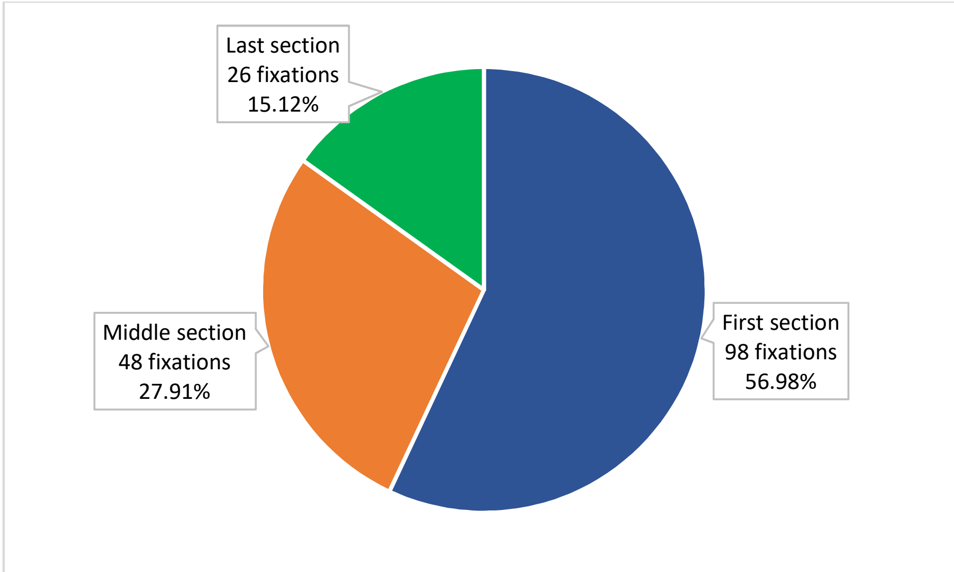


Figure 4.7. Non-expressive conductor video fixation count and percentages

Yet another pattern emerged when reviewing the Expressive conductor video. The middle section received the highest total - 128 fixations (49.81%). The count-off and first beat incorporated 84 fixations (32.68% of the total fixations), while the last measure garnered 45 fixations (17.51%). Figure 4.8 represents an increased percentage of fixations in the middle section when compared to the other conducting videos.

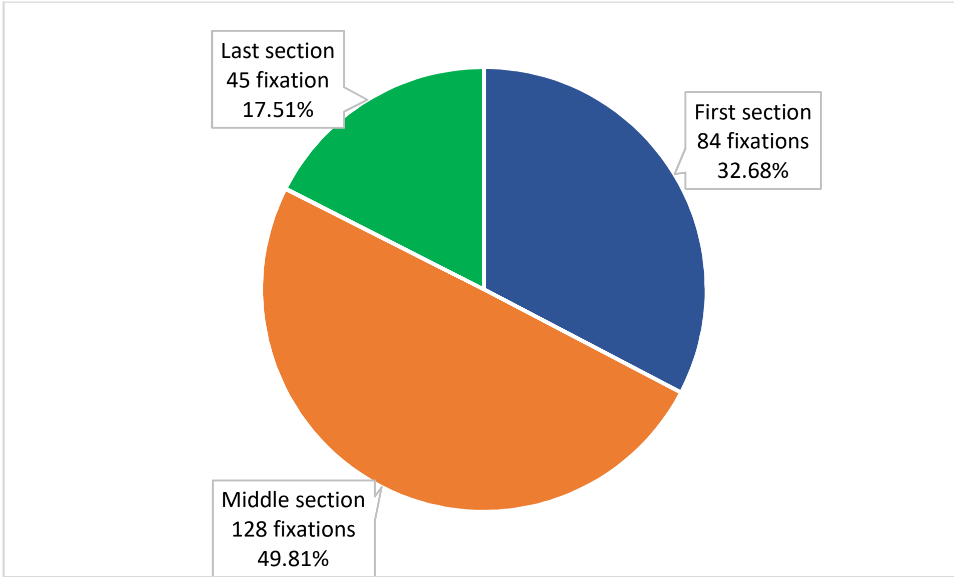


Figure 4.8. Expressive conductor video fixation counts and percentages

I used a repeated measures ANOVA to test for differences between the conductor videos' fixation count totals for the three sections. Neither the fixation count data from the count-off and first beat nor the final measure revealed a significant statistical difference. For the middle section, Mauchly's test of sphericity was significant, so the report shifted to Greenhouse-Geisser adjusted F-tests to control Type I errors. Results showed a significant main effect on the conductor videos, Greenhouse-Geisser Adjusted F (1.38, 28.88) = 7.60, p = .01, partial ω2 = .23. Pairwise comparisons revealed that the greater number of fixations on the Expressive conductor video was significantly higher than the fixation count on the Dot conductor video, p = .01. No significant differences occurred between the other videos.

Fixation durations for video sections

Figures 4.9, 4.10, and 4.11 present the same categories of data for each conductor video - mean fixation durations and the individual with the largest fixation duration along with the corresponding percentage of the time, plus the median fixation duration for each category. Comparing the data revealed that the total fixation duration, and therefore the percentages, increased from the Dot conductor video to the Non-expressive conductor video. The total and maximum fixation durations for the Expressive conductor video exceed the other conductor videos in all categories. The median fixation durations for the Dot and Non-expressive conductor videos were similar. However, the median fixation duration for Expressive conductor video was smaller for every category.

Dot conductor video	Entire video (47.5 seconds)	Count-off and first beat (3 seconds)	Middle section (42 seconds)	Last measure (2.5 seconds)
Mean total fixation duration [percent] (SD)	1.77 seconds [3.72%] (1.24)	1.13 seconds [37.73%] (0.70)	0.17 seconds [0.39%] (0.31)	0.47 seconds [18.95%] (0.67)
Longest fixation duration by an individual participant	4.45 seconds (9.37%)	2.45 seconds (81.67%)	1.24 seconds (2.95%)	1.78 seconds (71.20%)
Median fixation duration	0.19 seconds	0.21 seconds	0.19 seconds	0.19 seconds

Figure 4.9. Dot conductor video fixation durations and percentages

Non-expressive conductor video	Entire video (47.5 seconds)	Count-off and first beat (3 seconds)	Middle section (42 seconds)	Last measure (2.5 seconds)
Mean total fixation duration [percent] (SD)	2.41 seconds [5.08%] (1.63)	1.59 seconds [53.03%] (0.82)	0.44 seconds [1.06%] (0.70)	0.38 seconds [15.18%] (0.55)
Largest fixation duration by an individual participant	6.55 seconds (13.79%)	2.68 seconds (89.33%)	2.69 seconds (6.40%)	1.57 seconds (62.80%)
Median fixation duration	0.19 seconds	0.20 seconds	0.19 seconds	0.19 seconds

Figure 4.10. Non-expressive conductor video fixation durations and percentages

Expressive conductor video	Entire video (47.5 seconds)	Count-off and first beat (3 seconds)	Middle section (42 seconds)	Last measure (2.5 seconds)
Mean total fixation duration [percent] (SD)	3.03 seconds [6.39%] (2.34)	1.32 seconds [43.94%] (0.87)	1.10 seconds [2.61%] (1.47)	0.62 seconds [24.67%] (0.62)
Largest fixation duration by an individual participant	7.46 seconds (15.71%)	2.78 seconds (92.67%)	5.22 seconds (12.43%)	1.81 seconds (72.40%)
Median fixation duration	0.13 seconds	0.17 seconds	0.12 seconds	0.14 seconds

Figure 4.11. Expressive conductor video fixation durations and percentages

Figure 4.12 shows the fixation durations, and percentage thereof, the combined participants viewed each section of the Dot conductor video. Nearly two-thirds of the total time participants focused on this video occurred in the count-off and first beat. A comparison of the middle sections, color coded in orange, from Figure 4.12 to Figure 4.6 indicated that the total fixation counts equaled 15.54% yet the percentage of the total fixation duration equaled 9.34%. For the count-off and first beat, the percentage of the total duration, 63.91%, topped the percentage of the total fixation count, 56.89%. This means the average length of a fixation shrank during the middle section.

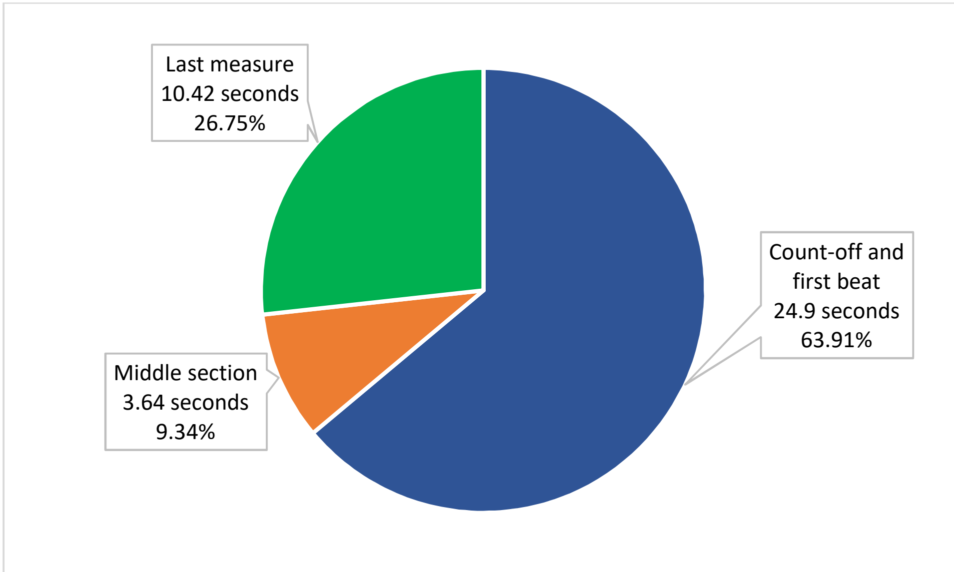


Figure 4.12. Combined fixation durations for all participants, Dot conductor video

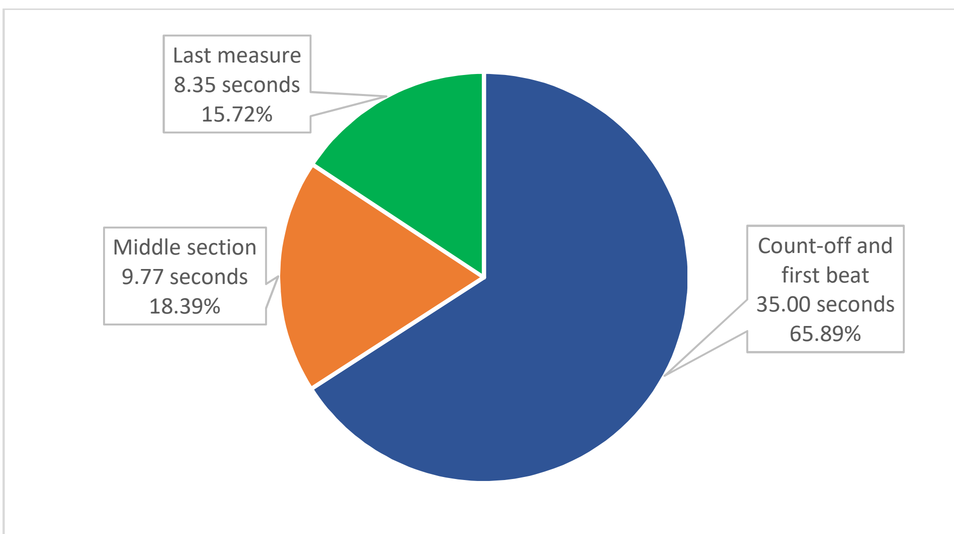


Figure 4.13. Combined fixation durations for all participants, Non-expressive conductor video

The same comparisons of data for the Non-expressive conductor video demonstrated the same trends as the Dot conductor video. The average fixation duration shrank after the count-off and first beat. An evaluation of the middle sections from Figure 4.13 and Figure 4.7 indicated that even though the total number of fixations equaled 27.91%, the percentage of the total viewing duration shrank to 18.39%. The percentage of the total duration in the count-off and first beat, 65.89%, surpassed the percentage of the total fixation count, 58.11%.

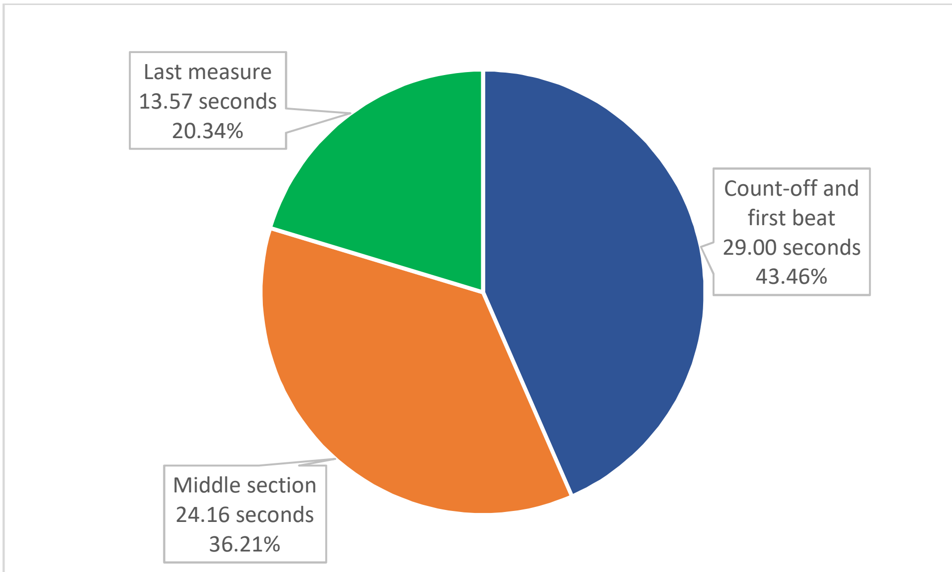


Figure 4.14. Combined fixation durations for all participants, Expressive conductor video

The assessment of the data for the Expressive conductor video exhibited a reduction of the average fixation duration after the count-off and first beat. The comparison of the middle sections from Figure 4.14 and Figure 4.8 indicated that even though the total number of fixations equaled 49.81%, the percentage of the total viewing duration shrank to 36.21%. During the count-off and first beat, the percentage of the total fixation duration, 43.46%, exceeded the percentage of the total fixation count, 32.68%.

Separate repeated measures ANOVA tests analyzed differences in fixation durations in each delineations of the videos. When comparing the fixation durations during the count-off and first beat, Mauchly's test of sphericity was not significant, $F(2, 42) = 6.82$, $p = .003$. A pairwise comparison showed the Non-Expressive Conductor Video had a significantly higher fixation duration than the Dot conductor video, $p = .001$. In testing fixation durations recorded in the middle section, Mauchly's test of sphericity was significant, so the report shifted to Greenhouse-Geisser adjusted F-tests to control Type I error. Results showed a significant main effect of Conductor Video, Greenhouse-Geisser Adjusted $F(1.47, 30.96) = 8.42$, $p = .003$, partial $\omega 2 = .25$. A pairwise comparison showed the fixation duration from the Expressive conductor video was significantly higher than the Dot conductor video, $p = .007$. Results from a repeated measures ANOVA test illuminated no significant difference between any of the fixation durations for the last measure.

Analysis of fixation counts for the middle section

The contrast in fixation counts during the middle section of the conductor videos needed further exploration. To facilitate this investigation, I generated a report through the Tobii Studio Pro software concerning the fixation counts for each measure from measure two to twenty-one. Figure 4.15 graphs the results of this report.

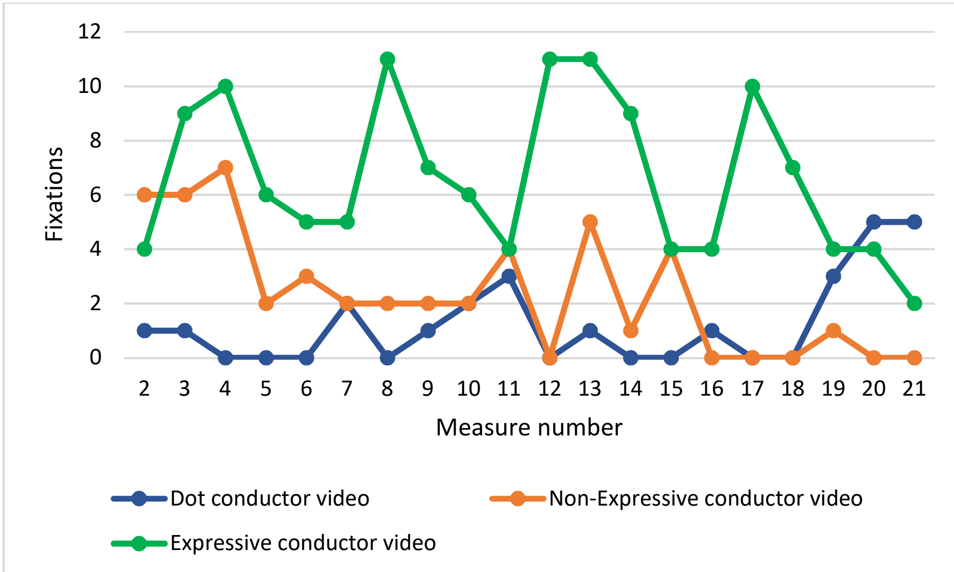


Figure 4.15. Fixation count per measure by conductor video

No participant fixated on the Dot conductor video in nine of the middle twenty measures, shown in Figure 4.15 by the blue line. In the corresponding span, participants did not fixate on the conductor during six measures of the Non-Expressive video, as portraited by the orange line. Interestingly, every measure of the Expressive conductor video, displayed in green, received at least one fixation. No predictable patterns appeared when viewing the graph. In seventeen of the featured twenty measures, the fixation count for the Expressive conductor video exceeded the other videos.

Fixations on different parts of the conductor's body

The Tobii Studio Pro software allowed for the creation of separate areas of interest. The specific areas created for this study encompassed the face and the torso. As reported in Figure 4.16, participants registered 33 fixations on the face and 105 fixations on the torso over the course of the Non-expressive conductor video, a ratio of approximately one to three.

Participants recorded a similar proportion of fixations for the Expressive conductor video, with 47 fixations on the face and 148 fixations on the torso.

	Expressive Conductor Video (Fixation Counts)		Non-Expressive Conductor Video (Fixation Counts)	
Participant	Face	Torso	Face	Torso
B1	2	15	2	10
B2	9	4	4	6
B3	0	4	0	3
C1	0	3	0	1
C3	6	14	2	6
C4	0	6	0	2
C5	6	24	6	7
C6	4	2	6	2
E1	7	7	3	7
F1	5	3	3	7
F2	1	7	0	9
H1	0	6	2	6
H2	0	9	0	6
H3	0	11	0	8
S1	2	9	0	9
S2	0	3	0	0
S3	0	1	4	1
TB1	1	6	0	5
Tr1	1	7	0	7
Tr3	3	0	1	1
Tr4	0	5	0	2
Tu1	0	2	0	0
Total	47	148	33	105

Figure 4.16. Fixation counts on conductor face and torso

Examination of the fixation durations, displayed in Figure 4.17, told a different story. Participants fixated on the conductor's torso for 30.51 seconds and on the face for 15.66 seconds during the Non-expressive conductor video, a two to one ratio. During the Expressive conductor video, the fixation times on the conductor's face and torso nearly matched, tallying

25.59 seconds and 28.04 seconds, respectively. Therefore, when participants chose to view the conductor's face, on average they held each fixation longer than upon the torso.

	Expressive Conductor Video (time in seconds)		Non-Expressive Conductor Video (time in seconds)	
Participant	Face	Torso	Face	Torso
B1	3.05	3.01	3.04	3.07
B2	6.02	0.43	0.92	2.44
B3	0	0.44	0	2.04
C1	0	0.31	0	0.08
C3	3.12	2.57	0.87	2.5
C4	0	0.53	0	0.28
C5	2.52	5.02	1.98	1.75
C6	3.44	0.77	4.07	0.24
E1	1.57	0.62	0.71	2.23
F1	0.51	0.73	0.34	1.62
F2	0.14	1.51	0	1.76
H1	0	1.97	0.25	1.23
H2	0	0.76	0	1.03
H3	0	2.29	0	2.57
S1	0.14	1.05	0	1.28
S2	0	0.26	0	0
S3	0	0.18	1.7	0.07
TB1	1.65	1.16	0	2.56
Tr1	0.79	3.66	0	2.95
Tr3	2.64	0	1.79	0.57
Tr4	0	0.46	0	0.22
Tu1	0	0.32	0	0
Total	25.59	28.05	15.67	30.49

Figure 4.17. Fixation durations on conductor face and torso

A cursory visual inspection of the gaze plot map of the human conductor videos suggested that more fixations occurred on the baton side, the right side, of the conductor than the left side. To confirm this assumption, the data for the torso was divided down the middle of the conductor's body. Overall, the baton side received 176 fixations for a duration of 45.52 seconds, while the left side garnered 77 fixations for 13.02 seconds. Since the baton occasionally crossed to the left side of the body, it is difficult to determine the significance of this number.

Misalignment and corrections

A practical purpose for a performer to view a conductor is to align with the beat pattern, and thereby synchronize with the rest of the ensemble. In listening to the performances during the study, I noticed some participants skipped notes at the end of a measure to play the first note of the next measure as the conductor gave the gesture for beat one. To further analyze this phenomenon, I created two additional categories: misalignments and corrections. Each time a participant did not sound the first note of each measure between the ictus of beat one and the ictus of beat two registered as a misalignment. Figure 4.18 presents the fixation counts from each person and the number of misaligned measures for each conductor video. I used a repeated measures ANOVA to test for differences between the conductor videos based on the numbers of misalignments. Results showed no significant statistical differences between any of the conductor video.

	Dot conductor video		Non-Expressive conductor video		Expressive conductor video		Total
Participant	Fixation count	Misaligned measures	Fixation count	Misaligned measures	Fixation count	Misaligned measures	Misaligned measures
B1	7	0	15	0	20	0	0
B2	19	12	10	15	14	9	36
B3	9	12	3	8	5	18	38
C1	6	7	3	0	2	0	7
C3	10	0	8	0	19	0	0
C4	2	0	2	0	5	3	3
C5	10	0	14	0	30	0	0
C6	8	0	8	0	7	5	5
E1	3	2	10	4	15	6	12
F1	7	1	15	0	12	1	2
F2	2	0	13	0	15	0	0
H1	7	13	8	13	5	14	40
H2	4	20	6	17	9	15	52
H3	10	8	8	0	10	0	8
S1	8	0	10	3	12	0	3
S2	1	6	2	6	4	12	24
S3	6	11	6	11	2	12	34
TB1	5	0	5	0	6	4	4
Tr1	8	2	7	0	8	0	2
Tr3	5	0	5	0	7	0	0
Tr4	0	0	5	0	13	13	13
Tu1	4	0	0	0	3	0	0
Total	141	94	163	77	223	112	283

Figure 4.18. Fixation counts and misaligned measure counts for all conductor videos

Since each person played the solo with three different conductor videos, the total number of measures for each participant to align with a conductor was 66. Given twenty-two participants, the total measures performed throughout the study equaled 1,452. For the entire study, the combined participants registered 283 total misaligned measures (19.49% of the total measures). During the Dot conductor video, participants registered 94 misaligned measures (19.24%). The misaligned measures in the Non-expressive conductor video added up to 77 (15.91% of the measures). Participants enumerated 112 misaligned measures (23.14%)

throughout the Expressive conductor video. Only one participant did not start the first measure at the correct time, which happened during the Expressive conductor video.

As can be seen in Figure 4.19, as the number of misalignments grew, illustrated by the blue line, the number of fixations jumped sporadically, depicted by the orange line. A Pearson product-moment correlation coefficient was calculated to determine if any relationship exists between the number of misalignments and the fixation counts. There was a weak negative correlation between the variables, $r = -0.20$. A separate Pearson product-moment correlation coefficient was calculated to determine if any relationship exists between the number of misalignments and the total fixation durations by each participant. There was a weak negative correlation between the variables, $r = -0.22$. For the participants in this study, increasing the number and duration of fixations led to a slight reduction in the number of misalignments.

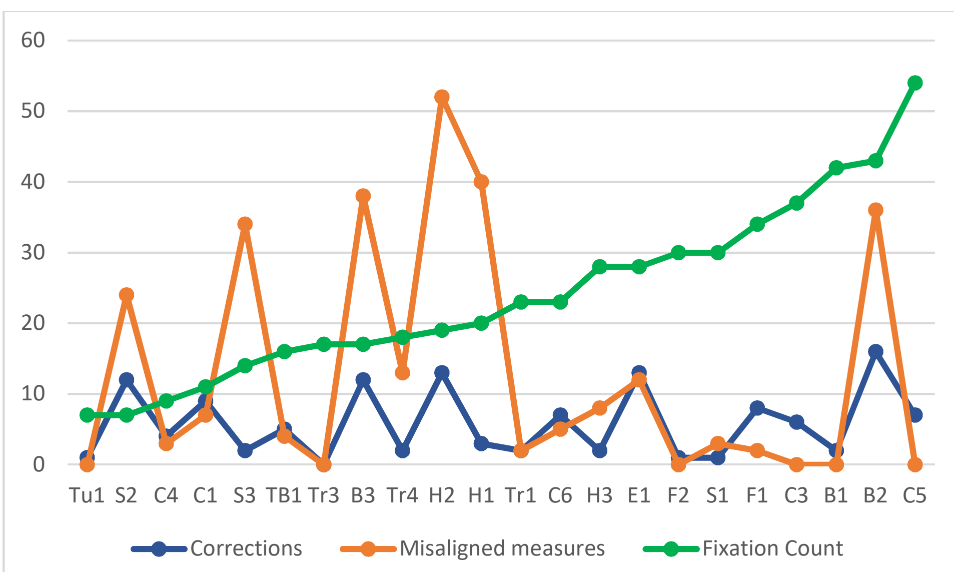


Figure 4.19. Comparison of misaligned measures to fixation counts

During this study, several participants attempted to realign their performance with the conductor's beat one gesture. This action gained the label "correction." Corrections encompassed the instances when a participant adjusted their place in the music to try to align with the conductor. Correction methods included skipping notes, pausing before playing the next note, severely rushing the tempo, lengthening note durations, and shortening note durations. Figure 4.20 lists the corrections by each participant for each video. Figure 4.20 also displays that one person made zero corrections when combining all three conductor videos.

Participant	Dot conductor video	Non-expressive conductor video	Expressive conductor video	Total corrections per Participant
B1	1	0	1	2
B2	7	5	4	16
B3	4	4	4	12
C1	4	3	2	9
C3	4	1	1	6
C4	2	1	1	4
C5	2	3	2	7
C6	2	1	4	7
E1	4	3	6	13
F1	2	2	4	8
F2	0	0	1	1
H1	1	1	1	3
H2	7	3	3	13
H3	1	0	1	2
S1	0	1	0	1
S2	3	4	5	12
S3	1	1	0	2
TB1	2	1	2	5
Tr1	1	1	0	2
Tr3	0	0	0	0
Tr4	1	0	1	2
Tu1	0	1	0	1
Total	49	36	43	128

Figure 4.20. Corrections per participant per conductor video

A Pearson product-moment correlation coefficient was calculated to determine if any relationship exists between the number of corrections and the fixation counts. There was a weak correlation between the variables, $r = .12$. A separate Pearson product-moment correlation coefficient was calculated to determine if any relationship exists between the number of corrections and the total fixation durations by each participant. There was a weak correlation between the variables, $r = .03$. These results both suggest that though the nature of the relationship is weak, participants who fixate on the conductor more are able to make more corrections.

Chapter 5: Discussion

The purpose of this study was to investigate the question "What are the visual patterns of woodwind and brass players on a conductor while reading music and playing an instrument?" To answer this inquiry, eye tracking technology collected data about instrumentalists' fixations while playing and reading music with a conductor. Several other questions stemmed from the initial query. Does the expressiveness of a conductor factor into the visual fixation patterns on a conductor by an instrumentalist? Does the complexity of the music affect visual patterns on a conductor? Could a conductor be replaced by an electronic facsimile? Does a participant need to foveally view a conductor to align the first note in each measure with the conductor's gesture? Do the fixation counts and durations increase with more accurate alignment?

Fixation duration on the conductor

The foundational query that inspired this study was a yes or no question, "Do members of an ensemble watch a conductor?" The definition of watch used in this study was based off of Findley and Gilchrist (2003) theory of active areas of attention; watch means to put purposeful attention on an area of the field of vision to gather information. A review of the data in my study showed all participants responded in some way to a conductor's gestures when they performed at least part of the music. Every participant directly fixated upon at least one conductor at some point during the three conductor videos. Since the participants did not fixate on the screen when a conductor was not present during the playing of the solo, I believe the fixations on the conductor were intentional acts and attempts to gather information. This affirms the notion that the participants watched a conductor.

As previously shown in Figures 4.9, 4.10, and 4.11, the mean total fixation durations amounted to small percentages of each video. The longest mean fixation duration occurred during the Expressive conductor video, lasting for 3.03 seconds out of 47.5 seconds. This equates to participants fixating on the conductor for 6.38% of the total time. The disparity between the mean fixation durations and the total video length is depicted in Figure 5.1. The data showed that participants viewed the Expressive conductor video for a significantly longer time than the Dot conductor video. Though statistical analysis found no significance differences, these participants did view the Non-expressive conductor video a longer duration than the Dot conductor video. The participants in this study preferred to fixate on the human conductors over the electronic facsimile. Figure 5.1 also suggests how much time it took for participants to gather detailed information through the fovea compared to the overall length of the video.

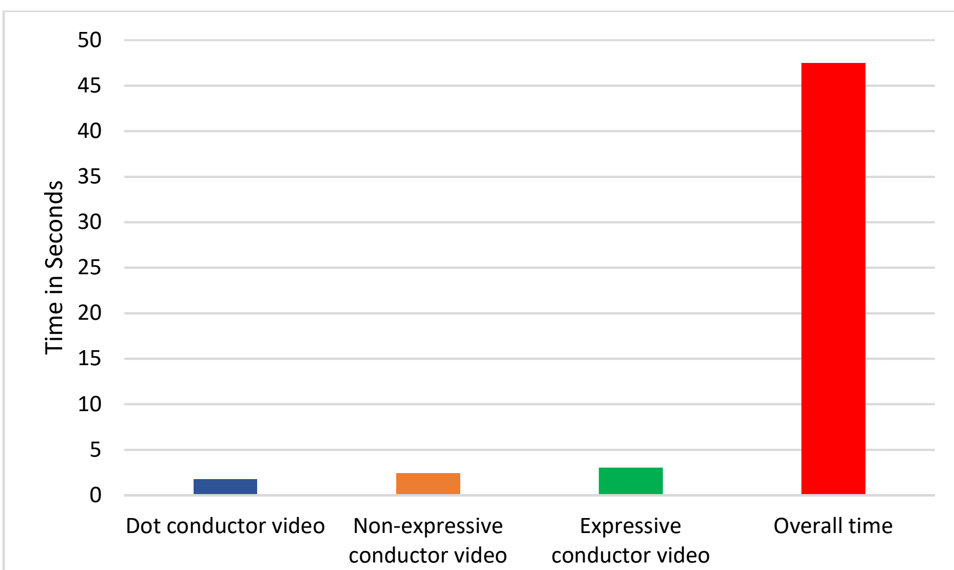


Figure 5.1. Fixation duration comparison of mean fixation durations to the total video length

Chapter 2 of this document cited factors affecting visual fixation patterns and mental processing when reading music. It is possible that the primary factor in determining the location of the participants' visual fixations was the difficulty of the music. The Expressive conductor exhibited expressive gestures, maintained eye contact, and provided an information rich environment. Though previous research suggests all these factors increase the likelihood of eye contact with a conductor, the participants only registered a mean fixation duration of 3.03 seconds, with the count-off and first beat accounting for over forty five percent of the total time. Seemingly, the participants emphasized fixating on the music over fixating on the conductor.

Several factors could explain this action. Since the participants had limited experience with the solo, they may have prioritized using the fovea to pick up the fine details of the printed music. In addition to that, the complexity of the music may have caused longer fixations allowing less time to fixate on the conductor. The participants may have used peripheral vision to gather all the information they felt they needed from the conductor. A final conjecture is that the participants trusted an internal construct of tempo and did not feel the need to fixate on the conductor.

Potential effects of the musical elements

Several factors potentially caused the disparity between Fredrickson's (1994) result of 28% of the time spent viewing the conductor and my study's result of 6.38%. Gudmondsdottir (2010) found that musicians with more experience could interpret more information in a shorter time. To attempt to match Fredrickson's study and control for experience bias, I selected participants from auditioned university bands. This is the only information available to compare the respective participants' experience levels.

The difference in complexity between the music incorporated in each study explain the variance in the percentage of time participants fixated on a conductor is. Wurtz, Mueri and Wiesendanger (2009) ascertained that the structure of the music mitigated experience and all musicians reduced their saccade lengths when reading a more complex piece of music. The primary discrepancy between the compositions used in my study and Fredrickson's study was the rhythmic structure. Fredrickson selected music with "rhythmic and technical simplicity" in mind and did not to use a subdivision smaller than quarter notes. My study's solo contained numerous eighth and sixteenth notes along with a few syncopated rhythms. Additionally, I incorporated few recognizable patterns and more unexpected notes in the form of accidentals, which Penttinen, Huovinen, and Ylitalo (2015) determined should extend the fixation durations of all musicians, regardless of experience. The complexity of the music in my study possibly caused the participants to fixate on the music longer, diminishing the available time to look at a conductor. Therefore, a plausible conjecture is the complexity of the music is inversely proportional to the time a musician fixates the conductor.

The actions of the conductor in Fredrickson's study may hold the key for explaining the increased fixations by participants. Since the video of the conductor used in Fredrickson's study is not available for analyzation, I can only offer conjectures. Possibly Fredrickson's conductor displayed even more facial characteristics, expressive gestures, and motion that drew more fixations from the participants. The participants may have been trained to fixate on that particular conductor prior to participant in that study.

Rayner (2009) reiterated the mean fixation duration when reading is 225 to 250 milliseconds. In the same article, he pointed out the mean fixation duration when freely viewing a scene is “close to 300 milliseconds” (p. 1476). Several factors affect fixation durations, especially when a specific task is involved. Henderson & Ferreira (1990) categorized the fixations that occur into two categories, “searching” fixations and "processing" fixations. A searching fixation is typically shorter, often less than 100 milliseconds. A processing fixation is longer, allowing an individual to process what they are viewing. As reported by Walshe and Nuthmann (2014), fixation duration is asymmetrical, meaning unexpected aspects in a scene lengthen a fixation duration, but expected aspects do not shorten the duration.

It is possible that musicians read a conductor’s gestures in the way they read music. When gestures are part of an expected “vocabulary” the mean fixation duration should be in the 225 to 250 millisecond range. For my study, participants recorded median fixation durations of 190-milliseconds for both the Dot and the Non-expressive conductor videos; the participants reduced their median fixation time to 130-milliseconds during the Expressive conductor video. One explanation for the short fixation durations in my study is participants more often searched for information than lingered to process it. Conversely, the participants may have not needed to search but instead quickly found the information they sought and transitioned to the next fixation point.

The Expressive conductor video included unexpected material which other research established should lengthen fixation durations, yet I cannot find evidence to support this. Possibly, the participants spent more time searching the Expressive conductor video for information than processing the image, which would bring the median time down. Since the No conductor instructional slide did not receive any fixations, the participants probably did not randomly shift their gaze to the conductor videos.

Kinsler and Carpenter (1995) advocated for a model of what happens in the brain that offers another plausible explanation as to why the median fixation duration decreased for the Expressive conductor video. To summarize their model, information is taken in by an encoder, such as the eye, moved to a processor, the visual cortex, assimilated as the visual cortex buffers this information in conjunction with other processors and previous memories, and then executed in the form of music performance. As the buffer clears, more material is taken from the processor. The processor then gathers the next round of information from the encoder, allowing the encoder to accept more information. The process is continuous.

The primary difference between conductor videos was the saturation of expressive information. Applying Kinsler and Carpenter's model to my study, one could posit the information rich environment created by additional gestures in the Expressive conductor video caused the encoder, processor, and buffer to fill with information in a shorter duration of time. Since the tempo compelled the participant to perform the music at a specific rate, the buffer, processor, and encoder needed to work faster. Therefore, the fixation duration decreased while the fixation counts increased.

Viewing the conductor during different video segments

To further analyze the data, I divided each conductor video into three sections: the count-off and first beat, the middle section, and the last measure. As seen in Figure 4.5, the sections represent the following percentages of the video: the count-off and first beat were

6.32%, the middle section encompassed 88.42%, and the last measure totaled 5.26%. However, the fixation durations and counts for each section do not reflect these proportions.

The longest fixation duration for every conductor video occurred during the count-off and first beat. During this frame, the mean fixation duration out of the total time ranged from 37.73% (Dot conductor video) to the significantly higher 53.03% (Non-Expressive conductor video). Figure 5.2, which displays mean fixation durations for all three conductor videos, acutely portrays how participants fixated more during the count-off and first beat over any other segment.

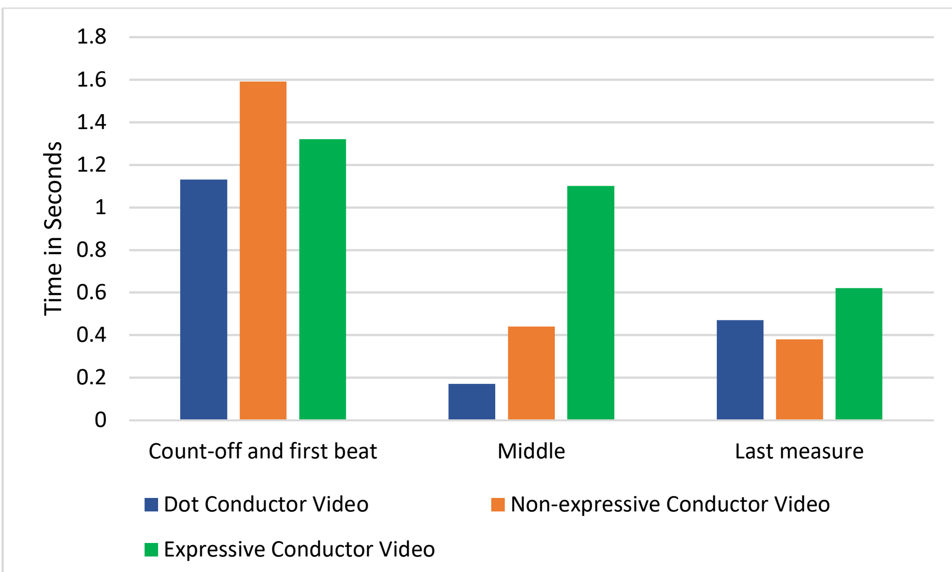


Figure 5.2. Mean fixation duration comparison for segments of conductor videos

Figure 5.3 illustrates the variability amongst participants' fixation counts through the three sections of the conductor videos. The count-off and first beat garnered the highest total fixation count, though a with-in subjects repeated measures ANOVA determined no significant interactions between any of the individual conductor videos. The tally of 268 fixations when combining the three videos in this three second window represents 46.45% of the entire 577 fixations.

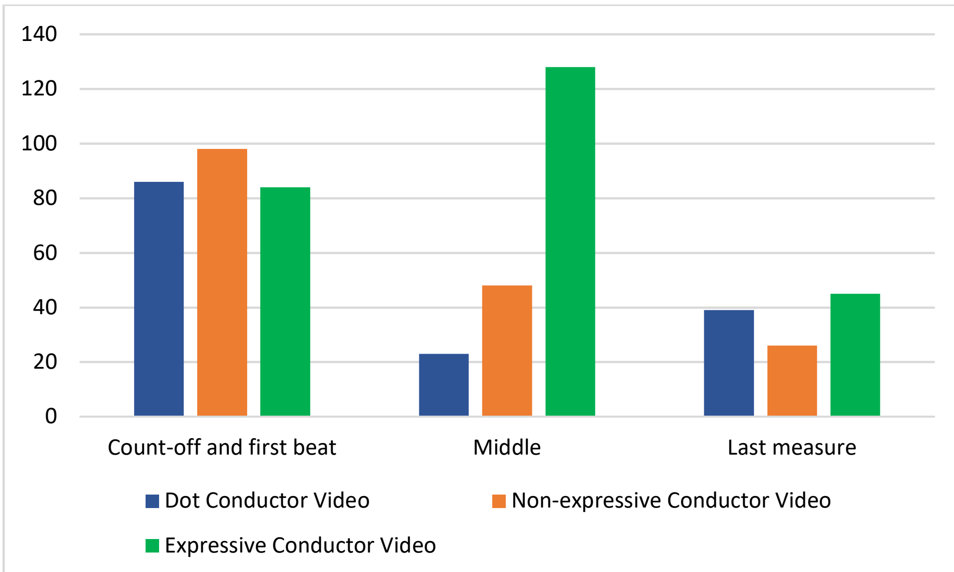


Figure 5.3. Total fixation count comparison for conductor videos

The longer fixation durations during the count-off and first beat suggest the participants emphasized beginning their performance at the moment of conductor initiation over incorporating other gestures in the subsequent measures. Another possible factor affecting the relatively high mean percentages of fixation durations may include the fact that participants did not perform during the count-off, nor did another informational source compete for attention. The simplicity of the music in the first measure might have allowed more time for the participants to fixate on the conductor.

The participants' previous ensemble training may explain the high fixation counts during the count-off and first measure. Some conductors begin music silently, meaning the ensemble members solely use visual cues to begin performances. Therefore, fovealy viewing the conductor before the music starts is a learned habit. Once the ensemble begins, these participants might listen to other ensemble members instead of watching the conductor. Even though my study removed all sound, the habit might remain, which may explain why the fixations dropped during the middle section of the videos.

Noticeable distinctions between the fixation counts during the middle sections of the conductor videos are also depicted in Figure 5.3. The 128 fixations on the Expressive conductor video are more than the fixation counts of the Dot (23) and Non-expressive (48) conductor videos combined. This data point supports House's (1998) finding that "expressive conducting contributes to more favorable performer attitudes toward the conductor." Meals (2018, p. 112) pointed out the results of his study "suggest that the quality of the expressive quality of the conductor's gesture may play a role in evaluations of efficacy. The one conductor whose gestures were noted as being more expressive than other conductors was rated higher across all offsets and in both fast and slow conditions." Possibly the information rich environment created in the Expressive conductor video, including independent hand gestures, varied facial expressions, and expressive conducting gestures, were essential to encouraging ensemble members to fixate on the director. Put another way, the preference for expressive conductors reported in other research (Morrison et al., 2009; Morrison et al., 2014; Morrison & Selvey, 2014; Price & Mann, 2009; Napoles, 2013) may manifest itself in the form of more fixations and longer fixation durations on the Expressive conductor video in my study.

An unexpected statistic is the time participants spent, or rather did not spend, fixating on the conductor videos during the middle segment. Though participants viewed the human conductor videos more than the Dot conductor video, they fixated on the middle section of the

Non-Expressive conductor video for a mean duration of just over one percent of the total time, 0.44 out of 42 seconds. The mean fixation duration more than doubled to 1.10 seconds during the corresponding section of the Expressive conductor video, but still only amounted to only 2.61% of the segment. There is a drastic difference between the mean times of each conductor video and the overall length of the segment from the middle section. The difference between the videos was the conductor. Therefore, I conclude that the actions of an expressive conductor will cause a higher fixation counts and longer fixation durations by ensemble members.

Misalignments, corrections, and the implied use of peripheral vision

The initial design of this study did not include tabulating misalignments or corrections. However, I observed several participants making corrections to apparently align with the conductor's gestures. The low percentage of fixation duration on the conductors did not seem to explain how the participants made corrections. An explanation is that participants relied heavily on their peripheral vision to maintain or regain alignment with the conductor.

As advocated by Findlay and Gilchrist (2003), a person can actively place attention in any portion of the visual field, they do not need to shift the fovea to intentionally gather information. The lack of direct focus on the Dot conductor video corroborates that theory. No participant fixated on the Dot conductor video in nine of the middle twenty measures (see Figure 4.15). Another five measures received only one fixation. Yet, participants aligned the playing of the first note with the beat one conducting gesture for over 80% of the measures in this video. The participants presumably absorbed information through peripheral vision to guide their alignment with the conductor's gestures.

Additional data concerning fixation counts per measure, fixation durations, misalignments, and corrections also support Findlay and Gilchrist's concept of actively incorporating information appearing in all visual fields. During the Dot conductor video, the participants garnered the fewest total fixations and shortest fixation durations. The total number of misaligned measures for all participants, 94, was not statistically different from either of the other videos. The number of corrections during the Dot conductor video, 49, was the highest total. Plausibly, the participants incorporated information gathered through their peripheral vision to make the corrections.

As previously discussed, I created three conductor videos to assess participant reactions to different types of conductor performances, which is in essence a virtual environment. Other research looked in to how musicians react in virtual environments. Williamon, Aufegger, and Eiholzer (2014) created two virtual scenarios of the stressful situations that student musicians experience, a recital and a jury. Each simulation contained all the pre- and post-performance protocols, along with life-sized virtual humans found in each venue. The participants rated the experience as realistic and reported the scenarios caused higher reported anxiety levels and heart rates. The researchers gathered electrocardiographic data from a subset of the participants during both the virtual experience and an actual audition in front of live judges. A comparison of these results showed suggested the virtual experience closely approximated the real audition. Studies by Orman (2004) and Aufegger, Perkins, Wasley, and Williamon (2017) have reported similar results.

Aspects of my study support the idea that the participants performed in the virtual experience of the conductor videos in a manner similar to how they might respond to a live conductor. Anecdotally, in my observations of live rehearsals, I witnessed musicians correct their place in the music to align with the conductor. When presented with the No conductor instructional slide, only one participant in this study skipped notes. If the participants did not care about aligning with the conductor, the number of corrections during the conductor videos would be similar to the No conductor instructional slide.

The answer to the question “Could a human conductor be replaced by an electronic facsimile?” is yes, assuming the purpose of a conductor is only to start, stop, and to keep tempo for the ensemble. Sidoti (1990) and Thompson (2012) support the finding that musicians are more likely to perform musical marking when a conductor performs an analogous gesture. While my study suggests players would rarely fixate on the electronic facsimile (in this case the Dot conductor video) after the count-off and first beat, I did not include evaluations of the expressiveness or accuracy (beyond the timing of the first note of each measure) during each participant’s performance. The data from this study indicate that the participants fovealy focused on the human conductor videos for a longer duration and with a higher fixation counts but cannot support conclusions concerning the relationship between expressive conducting and expressive performance.

The Non-Expressive conductor video displayed some information not included in the Dot conductor video. The major alteration was the presence of a human who made eye contact with the camera. The use of the left-hand added motion. This important alteration in the amount information being conveyed might explain the differences. In every measure except for the final one, participants registered longer fixation durations and higher fixation counts during the Non-Expressive conductor video than the Dot conductor video.

The Expressive conductor video presented the most information rich environment. The Expressive conductor video featured a conductor making eye contact while changing facial expressions. The left hand moved independently of the right hand. The pattern size changed, as did the style indicated. This video received the highest total fixation count and longest fixation duration. Based on this data, the expressiveness of a conductor appears to lead to increased fixations on the conductor.

The Expressive conductor video had the greatest number of misaligned measures. This video also recorded the smallest mean fixation duration. This conceivably illuminates processing complications experienced by participants as they tried to incorporate the additional material. Seemingly, fewer participants buffered the information at the appropriate time, resulting in a greater number of misalignments.

Wong and Gauthier (2012) found evidence that music-reading experts acquired visual skills resulting in higher music-reading fluency. Expanding from their finding, it is plausible fixation counts and durations are inversely proportional to misalignments. A Pearson product-moment correlation coefficient calculated a weak negative correlation. As participants increased the number and duration of fixations, the number of misalignments decreased slightly. Though it is difficult to see a pattern in Figure 5.4 due to the wide variance in the number of misalignments, signified by the orange line, the tendency is to have fewer misalignments as fixations increase. Highlighting the six participants with zero misalignments, the total fixation counts range from the highest number of fixations (participant C5 with 54) to the lowest (participant Tu1 with 7). At the opposite end of the spectrum, participant B2 recorded the second highest fixation count, 43, and yet misaligned more than half of the total measures, 36.

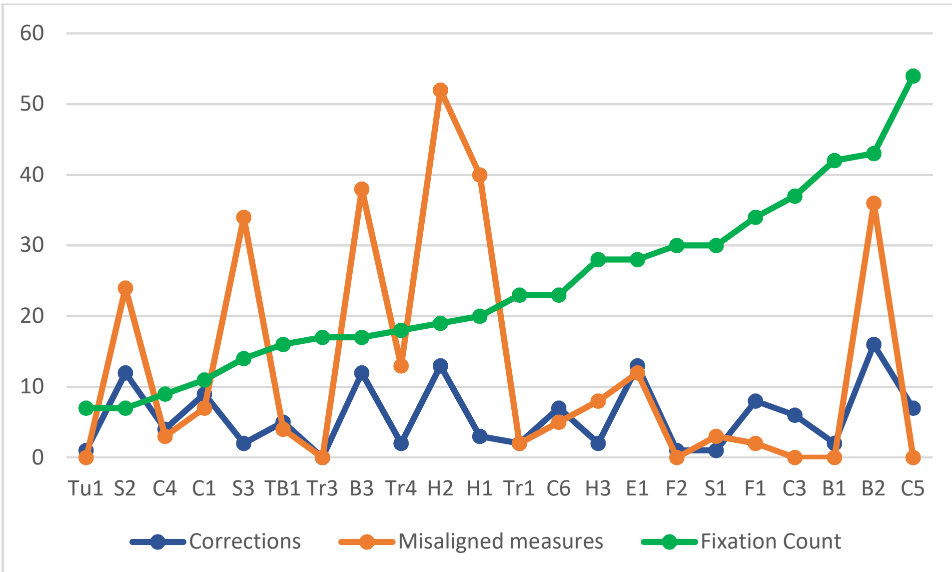


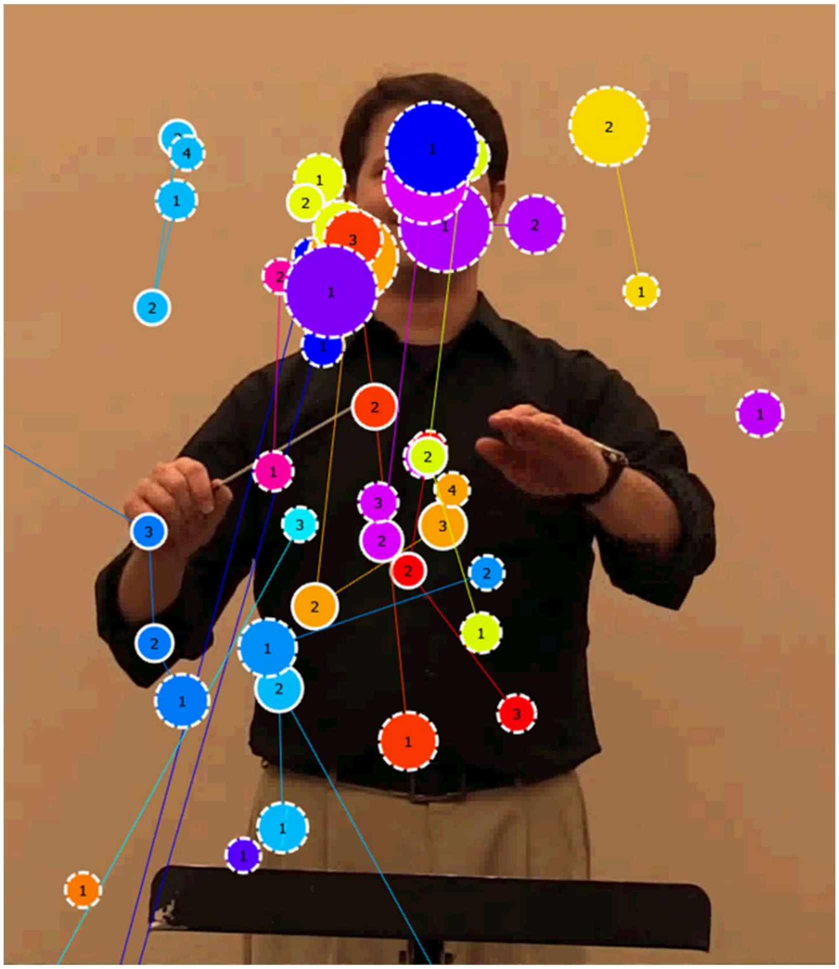
Figure 5.4. Comparison of Fixation Counts, Misalignments, and Corrections

Included on Figure 5.4 are the number of corrections, denoted by the blue line. A Pearson product-moment correlation coefficient calculated weak positive correlation between the number of corrections. As with misalignment, the pattern between corrections and fixation counts is hard to discern. A specific comparison of participants Tu1 and S2 highlights the disparity. Each participant accumulated 7 total fixations. Even so, participant Tu1 registered the second fewest corrections, one, while participant S2 recorded twelve corrections, among the highest total. Participants B1 and B2 enumerated high fixation counts, but their numbers of corrections demonstrate a wide discrepancy.

One of the early questions concerned musicians' ability to respond to a conductor's gestures while playing the music. My study indicates that increasing the fixations on the conductor with the fovea has a marginal effect on decreasing misalignments and increasing the chances a musician will be able to make a correction so they can align with the conductor's beat one. Peripheral vision also seems to be at least as important a method for participants to view the conductor. Regardless of the field of vision, the respondents' efforts, in the form of making corrections to align with the start of each measure confirm that musicians respond to the gestures of a conductor.

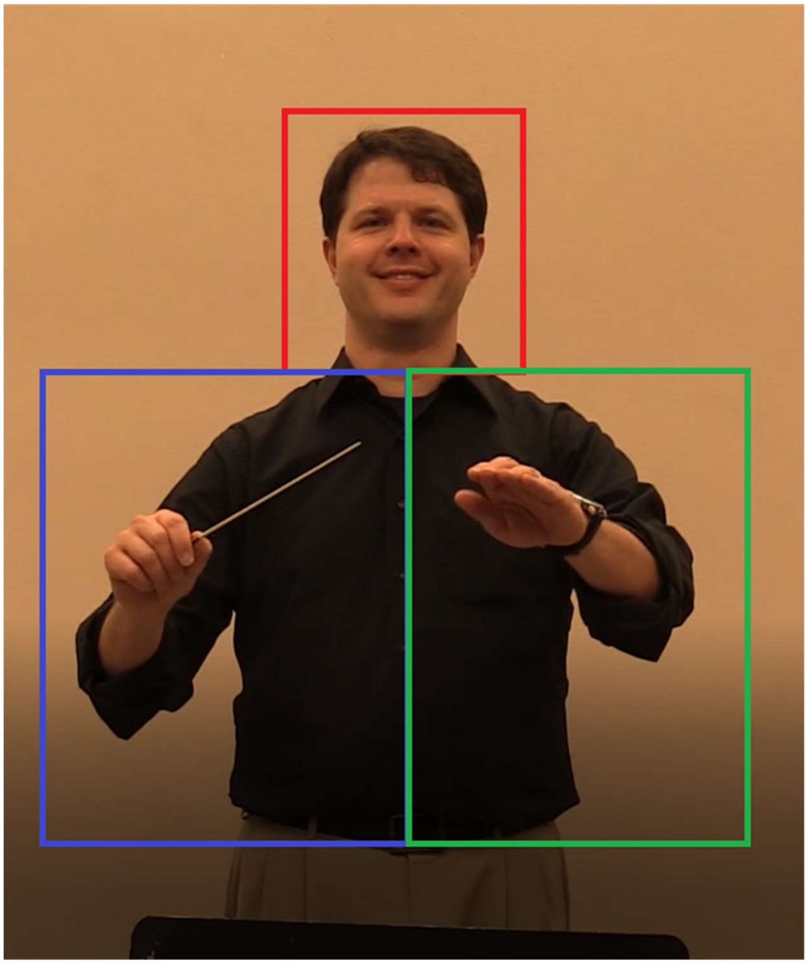
Where the participants fixated

Picture 5.1 shows a single frame of the sliding gaze plot map captured from the Expressive conductor video. As with the gaze plot maps of reading words and music, the size of the circle increases with the length of the fixation. The lines between circles represent saccades. Any line leading out of the frame signifies the software tracked the fovea as a participant's gaze either entered or left the screen. Each color indicates a different participant. The numbers in each circle represent the order of fixations.



Picture 5.1. Locations of fixations on Expressive conductor video

Viewing the gaze plot maps in motion for the human conductor videos suggested that participants fixated primarily on two areas: the face and the right side of the torso. Picture 5.2 depicts the areas of interest created to explore this observation: the face (red box), the right side of the conductor's body which encompassed the primary location for the baton (blue box), the left side of the body (green box), and the entire torso (combined blue and green boxes). Each of these areas encompassed the space through which all movement occurred. The baton crossed the center line of the body more throughout the Expressive conductor video than during the Non-expressive conductor video.



Picture 5.2. Areas of interest

An investigation of the data from each area of interest revealed the participants fixated on the conductor’s torso more times and for a longer duration than the face, with the right side receiving higher totals than the left side. The right side received a significantly higher number of fixations than the face. Since the baton occasionally traveled across all parts of the body, this result could be skewed. A plausible conclusion is the baton may attract the attention. The baton moved more and provided more information, including tempo, style, and dynamics, than the other parts of the conductor’s body.

A plausible explanation for participant's preference to view the torso more than the face could be the type of information the participants prioritized. Wollner (2007) found the conductor's arms gather more audience attention due to continuous motion. In a separate study, Wollner (2008) reported participants' survey responses indicating the conductor's arms provided more general information, but the face delivered more expressive information. In my study, the facial expressions changed only a few times during the Expressive conductor video and did not change throughout the Non-expressive conductor video. The participants possibly determined that they needed the information provided by the arms more than the expressive information in the Expressive conductor's face. Alternately, Calvo, Beltrán, and Fernández-Martín, (2014) found participants could categorize facial expressions shown in their peripheral vision. It could be possible the participants registered the changes in facial expressions without needed to fixate on the face.

While the torso had both a higher fixation count and fixation duration than the face, the average fixation duration on the face was longer. The fixation count ratio for both human conductor videos when comparing the torso to the face was approximately a three to one. Yet the fixation duration ratio when comparing the torso to the face for the Non-expressive conductor was two to one and the Expressive conductor approached a one to one ratio. This could indicate that the information displayed by the arms was processed faster. Alternatively, it could reflect a need to fixate on the face longer in order to process the facial expressions. It could be the participants placed greater emphasis on accurately processing the expressiveness of the face.

Implications

Conductors should understand that ensemble members can incorporate some information from the conductor's gestures even if they do not fixate on the conductor. Since peripheral vision seems to be the method for the intake of the majority of information, conductors must select gestures that ensemble members can see. In practical terms, those in the back of the room may not be able to detect small motions. Conductors should coordinate their gestures when moving their hands independently. Conflicting information may lead to unintended results. A musician using only peripheral vision will most likely see the motion of the baton, though they may also register the facial expression. A musician who fixates on the conductor more often locates the fixation on the right side of the body. If the arms are spread wide, the information it conveys may be lost.

The less expressive the conductor, the less ensemble members seem to fixate on the conductor. Adding the results of my study to the results published by Fredrickson (1994) and Byo and Lethco (2001), all the data point to instrumentalist's preference to view a conductor who uses expressive gestures. An additional benefit to increasing the fixations on the conductor is the weak correlation showing that misalignments will decrease and corrections will increase as the number of fixations increase.

Musicians may struggle to incorporate all of the conductor's gestures while playing. The greatest number of misalignments and corrections occurred during the Expressive conductor video. Increasing the experience level and knowledge base of the responding musicians may reduce these numbers. Some of the gestures may have not been within the participant's lexicon of knowledge, causing longer processing durations and sacrificing performance accuracy.

Conductors should teach their ensembles to recognize gestures, thereby building skill sets and enhancing performer's abilities to incorporate this additional information. Conductors should combine this training with practice on other skills. For instance, musicians can practice reading music, responding to the conductor, and executing expressive notation at the same time.

Conductors should be consistent when conducting the ensemble. Penttinen, Huovinen, and Ylitalo (2015) reported unexpected notes caused longer fixations, regardless of the responding musician's experience level. Wurtz, Mueri and Wiesendanger (2009) reached a similar conclusion, suggesting experience is mitigated by the structure of the music. If the conductor changes their actions every run-through, they could cause musicians to fixate on them for longer periods. While that may seem like a positive outcome, it may come at the expense of accuracy of performance, meaning performers spend more time processing and buffering the new gestures, leaving less time for the processing and buffering of other aspects of the performance. In my study, the participants only saw each video once, meaning they were sight-reading the conductor's gestures. The Expressive conductor contained the greatest number of expressive gestures and resulted in the largest number of misalignments.

The complexity of the music apparently affects ensemble members' propensity to fixate on the conductor. Experience can overcome the effects of complexity. Instead of handing out music and immediately playing through the piece, conductors should consider giving the ensemble time to practice it on their own. They could identify parts that relate to musician's previous experience or create exercises to practice the complicated parts. Any of these suggestions may lead to the ensemble incorporating more information from the conductor's gestures.

A conductor's garments or the background behind the podium may also affect the ability to discern conducting gestures, especially those made by the baton. As Bouma (1970) first described, people need a certain amount of critical spacing needed to recognize objects. If the conductor chose to wear clothing the same color as the baton or with an assertive pattern, the image becomes crowded. The garment provides too much relevant information making the baton harder to separate from the background. Compounding this problem may be a musician's use of peripheral vision, which lacks the ability to decipher fine details.

Limitations of this study

While the current study offered empirical data concerning fixations on a conductor, there are limitations. The technology used in this study recorded the visual patterns on videos of conductors. While this virtual scenario should produce results comparable to the responses to a live conductor, confirmation is needed. Also, the technology could not monitor the visual patterns on the music.

The participants were undergraduate students from three public universities in the Pacific Northwest. The results may not apply to other skill levels or locations. Neither the number of participants from individual instruments nor instrument family were large enough to be able to determine if difference exist between sub-groups. Not all instrument or instrument families were represented in this study. These results are based on a situation that did not include sound, aside from the participant's own instrument.

Participants only viewed the complete videos once, simultaneously executing the solo; in essence, they were sight-reading the conductor. Additionally, the participants did not have a relationship with the conductor, increasing the chances they may not have recognized all of the gestures.

Suggestions for further research

To eliminate a confounding factor, sound was intentionally removed from this study. A replication of this study with the addition of sound might reveal different visual. Methods to accomplish this are to perform the study during a live ensemble rehearsal (a model similar to Byo and Lethco's 2001 study), play a recording of an ensemble with parts different from the study participant's parts, play a recording identical to the part played by a participant, or a combination of all these approaches. The use of eye-tracking systems on multiple people simultaneously during a live rehearsal could incorporate visual patterns on both the conductor and the music, combining my study's fixation data with Byo and Lethco's system of "looks per measure" and furthering the knowledge about ensemble members visual patterns.

The conducting videos created for this study were intentionally recorded from the same distance and angle; however, not every ensemble member views the conductor at the identical angle. A follow-up study should be undertaken where a conductor is recorded from differing angles and distances. Comparisons should be made between the angles of the camera, as well as the location each participant sits relative to the conductor's podium.

Future research should address the question of transferability. Replications should include other age and experience parameters, a broader array of instruments, and other populations. A longitudinal study concerning the visual patterns through the rehearsal process might shed light on if or how changes visual patterns occur.

Another potential confounding factor purposely excluded from this study were musical markings, including dynamics and style. Byo and Lethco (2001, p. 21) stated “music characterized by frequent change was more conducive to watching the conductor than music of a static nature.” A replication of that study, using modern eye-tracking technology, would show the specific locations on either the conductor or the music participants viewed when music events occurred. Syncing a participant’s gaze plot map with their audio recording may enhance the study further.

My study shows a that participants fixated more on an expressive conductor. A follow-up study should explore if a type of expression garners more fixations than other types. For example, one entire video could be a smooth expressivo with a matching facial expression and the next could be bouncy with a complimentary facial expression. Both of these are expressive in nature but may not produce the same results.

A version of this study in which the participants receive training on the conducting gestures that will be incorporated may alter the visual fixation patterns. The goal of this suggested design would be to make sure the participants were not sight-reading the conductor’s gestures. A variation on this design would be to allow the participants to view the entire video without playing the music.

Conclusion

This study was one of the first to measure fixation counts and durations on a conductor, and as such it has created more questions than answers. The participants fixated on the conducting videos a small percentage of time, even when presented with a conductor who displayed numerous factors that contribute to higher viewing rates. The results provide data supporting conclusions that musicians prefer to view an expressive conductor. The study also suggests that the complexity of the music may affect the number and duration of fixations on the conductor. Finally, this study indicates peripheral vision factors prominently into the information incorporated by an instrumentalist.

References

Aufegger, L., Perkins, R., Wasley, D., & Williamon, A. (2017). Musicians' perceptions and experiences of using simulation training to develop performance skills. *Psychology of Music,* 45, 417-431.

Bebko, J., Uchikawa, M., Saida, K., & Ikeda, S. (1995). Learning to Recognize Patterns: Changes in the Visual Field with Familiarity. *Optical Review,* 1, 72-80.

Bevilacqua, A., Paas, F., & Krigbaum, G. (2016). Effects of motion in the far peripheral visual field on cognitive test performance and cognitive load. *Perceptual and Motor Skills, 122*, 452-469.

Bishop, L., & Goebl, W. (2015). When they listen and when they watch: Pianists' use of nonverbal audio and visual cues during duet performance. *Musicae Scientiae, 19*, 84-110.

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature, 226*, 177-178.

Byo, J., & Lethco, L. (2001). Student musicians' eye contact with the conductor: An exploratory investigation. *Contributions to Music Education, 28,* 21-35.

Calvo, M., Beltrán, D., & Fernández-Martín, A. (2014). Processing of facial expressions in peripheral vision: Neurophysiological evidence. *Biological Psychology, 100*, 60-70.

Calvo, M., & Beltrán, D. (2013). Recognition advantage of happy faces: Tracing the neurocognitive processes. *Neuropsychologia, 51*, 2051-2061.

Clayton, A. (1986) *Coordination between players in musical performance* (Unpublished doctoral dissertation). Edinburgh University, UK.

Cofer, R. (1998). Effects of conducting-gesture instruction on seventh-grade band students' performance response to conducting emblems. *Journal of Research in Music Education, 46*, 360-373.

Drieghe, D., Brysbaert, M., & Desmet, T. (2005). Parafoveal-on-foveal effects on eye movements in text reading: Does an extra space make a difference?. *Vision Research, 45*, 1693-1706.

Findlay, J., & Gilchrist, I. (2003). *Active vision: The psychology of looking and seeing*. Oxford: Oxford University Press.

Fredrickson, W. (1992). *An exploratory study of band musicians' eye contact and performance as influenced by loss of visual and/or aural stimuli* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses.

Fredrickson, W. (1994). Band musicians' performance and eye contact as influenced by loss of a visual and/or aural stimulus. *Journal of Research in Music Education, 42*, 306-317.

Goolsby, T. (1994). Profiles of processing: Eye movements during sightreading. *Music Perception*, 12, 97-123.

Gudmundsdottir, H. (2010). Advances in music reading research. *Music Education Research*, 12, 331-338.

Henderson, J., & Ferreira, F. (1990). Effects of fovea processing difficulty on the perceptual span in reading: implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 417-429.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Halszka, J., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford: Oxford University Press.

House, R. (1998). *Effects of Expressive and Nonexpressive Conducting on the Performance and Attitudes of Advanced Instrumentalists*, ProQuest Dissertations and Theses.

Johnson, C., Fredrickson, W., Achey, C., & Gentry, G. (2003). The effect of nonverbal elements of conducting on the overall evaluation of student and professional conductors*. Journal of Band Research*, 38, 64-77.

Kinsler, V., & Carpenter, R. (1995). Saccadic eye movements while reading music. *Vision Research*, *35*, 1447-1458.

Levy-Schoen, A. (1981). Flexible and/or rigid control of oculomotor scanning behavior. *Eye movements: Cognition and visual perception* (pp. 299-316). Hillsdale, NJ: Erlbaum.

Luck, G., & Nte, S. (2008). An investigation of conductors' temporal gestures and conductor-musician synchronization, and a first experiment. *Psychology of Music, 36,* 81-99.

MacKenzie, C., Vaneerd, D., Graham, E., Huron, D., & Wills, B. (1986). The effect of tonal structure on rhythm in piano performance. *Music Perception, 4,* 215-225.

Madell, J., & Hebert, S. (2008). Eye movements and music reading: Where do we look next? *Music Perception: An Interdisciplinary Journal, 26,* 157-170.

Morrison, S., Price, H., Smedley, E., & Meals, C. (2014). Conductor gestures influence evaluations of ensemble performance. *Frontiers in Psychology, 5*, 806.

Morrison, S., & Selvey, J. (2014). The effect of conductor expressivity on choral ensemble evaluation. *Bulletin for the Council of Research in Music Education, 199*, 7-18. doi: http://www.jstor.org/stable/10.5406/bulcouresmusedu.199.0007

Morrison, S., Price, H., Geiger, C., & Carnacchio, R. (2009). The effect of conductor expressivity on ensemble performance evaluation. *Journal of Research in Music Education, 57* (1). 37-49.

Napoles, J. (2013). The influences of presentation modes and conducting gestures on the perceptions of expressive choral performance of high school musicians attending a summer choral camp. *International Journal of Research in Music Education,* 31*,* 321- 330.

Niehorster, D., Cornelissen, T., Holmqvist, K., Hooge, I., & Hessels, R. (2017). What to expect from your remote eye-tracker when participants are unrestrained. *Behavior Research Methods.* https://doi.org/10.3758/s13428-017-0863-0

Orman, E. (2004). Effect of Virtual Reality Graded Exposure on Anxiety Levels of Performing Musicians: A Case Study. *Journal of Music Therapy,* 41, 70-78.

Palmer, C., & Krumhansl, C. (1987). Pitch and temporal contributions to musical phrase perception: Effects of harmony, performance timing, and familiarity. *Perception & Psychophysics, 41,* 505-518.

Pelli, D., & Tillman, K. (2008). The uncrowded window of object recognition. *Nature Neuroscience*, *11,* 1129-1135.

Penttinen, M., Huovinen, E., & Ylitalo, A. (2015). Reading ahead: Adult music students' eye movements in temporally controlled performances of a children's song. *International Journal of Music Education, 33*, 36-50.

Polanka, M. (1995). Research note: Factors affecting eye movements during the reading of short melodies. *Psychology of Music*, *23,* 177-183.

Price, H. & Mann, A. (2011). The effect of conductors on ensemble evaluations. *Bulletin of the Council for Research in Music Education,* 189*.* 57-72. doi: http://www.jstor.org/stable/10.5406/bulcouresmusedu.189.0057

Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*, 1457–1506. https://doi.org/10.1080/17470210902816461

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372-422.

Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology, 41,* 211-236.

Ryu, D., Abernethy, B., Mann, D., Poolton, J., & Enns, J. (2015). The contributions of central and peripheral vision to expertise in basketball: How blur helps to provide a clearer picture. *Journal of Experimental Psychology: Human Perception and Performance, 41*, 167-185.

Schwartz, S., Vuilleumier, P., Hutton, C., Maravita, A., Dolan, R., & Driver, J. (2005). Attentional load and sensory competition in human vision: Modulation of fMRI responses by load at fixation during task-irrelevant stimulation in the peripheral visual field. *Cerebral Cortex, 15*, 770-786.

Salis, D. (1980). Laterality effects with visual perception of musical chords and dot patterns. *Perception & Psychophysics, 28*, 284-92.

Sidoti, V. (1990). *The effects of expressive and nonexpressive conducting on the performance accuracy of selected expression markings by individual high school instrumentalists* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses.

Silvey, B. (2013). The Role of Conductor Facial Expression in Students' Evaluation of Ensemble Expressivity. *Journal of Research in Music Education, 60*, 419-429.

Sloboda, J. (1977). Phrase units as determinants of visual processing in music reading. *British Journal of Psychology, 68,* 117-24.

Sloboda, J. (1983). The communication of musical metre in piano performance. *Quarterly Journal of Experimental Psychology 35,* 377-96.

Sloboda, J. (1985). *The musical mind: The cognitive psychology of music.* Oxford, UK: Oxford University Press.

Thompson, J. (2012). *The effects of conducting-gesture instruction on high school string orchestra students' recognition of and playing response to common musical conducting emblems* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses.

Traschütz, A., Zinke, W., & Wegener, D. (2012). Speed change detection in foveal and peripheral vision. *Vision Research, 72*, 1-13.

Truitt, F., Clifton, C., Pollatsek, A., & Rayner, K. (1997). The perceptual span and the eye-hand span in sight reading music. *Visual Cognition, 4*, 143-161.

Walshe, R., & Nuthmann, A. (2014). Asymmetrical control of fixation durations in scene viewing. *Vision Research, 100*, 38-46.

Waters, A., Townsend, E., & Underwood, G. (1998). Expertise in musical sight reading: A study of pianists. *British Journal of Psychology, 89*, 123-149.

Waters, A., & Underwood, G. (1998). Eye movements in a simple music reading task: A study of experts and novice musicians. *Psychology of Music*, *26,* 46-60.

Williamon, A., Aufegger, L., & Eiholzer, H. (2014). Simulating and stimulating performance: Introducing distributed simulation to enhance musical learning and performance. *Frontiers in Psychology, 5*, 1-9.

Wong, Y., & Gauthier, K. (2012). Music-reading expertise alters visual spatial resolution for musical notation. *Psychonomic Bulletin & Review, 19*, 594-600.

Wurtz, P., Mueri, R., & Wiesendanger, M. (2009). Sight-reading of violinists: Eye movements anticipate the musical flow. *Experimental Brain Research, 194*, 445-450.

WÖllner, C. (2008). Which part of the conductor's body conveys most expressive information? A spatial occlusion approach. *Musicae Scientiae, 12*, 249-272.

WÖllner, C. (2007). *Zur Wahrnehmung des Ausdrucks beim Dirigieren. Eine experimentelle musikpsychologische Untersuchung* [The perception of expressiveness *in* conducting. An experimental music-psychological investigation]. Berlin: Lit-Verlag.

Appendix A: Consent Form

UNIVERSITY OF WASHINGTON

CONSENT FORM

Musician's Visual Fixations Study

Primary Researcher: Doug Morin
Contact Information: djmorin@uw.edu. 206-543-0824.
Advisors: Dr. Steve Morrison, Prof. Timothy Salzman

<u>Researchers' statement</u>

We are asking you to be in a research study. The purpose of this consent form is to give you the information you will need to help you decide whether to be in the study or not. Please read the form carefully. You may ask questions about the purpose of the research, what we would ask you to do, the possible risks and benefits, your rights as a volunteer, and anything else about the research or this form that is not clear. When we have answered all your questions, you can decide if you want to be in the study or not. This process is called "informed consent." We will give you a copy of this form for your records.

PURPOSE OF THE STUDY

This study investigates the visual fixations of a wind instrumentalist while reading music, playing an instrument, and a conductor is conducting.

STUDY PROCEDURES

The study will last no more than 30 minutes. Each participant should provide their own instrument. Music will be provided.

After a sight reading period, each person will play the music 4 times. During these performances, eye-tracking technology will monitor visual fixations. The technology uses infrared light. A video will be taken, but will not show a participant's face.

RISKS, STRESS, OR DISCOMFORT

Recordings will be destroyed after the data is analyzed. The identity of all participants will be kept separate from any data gathered. The audio/video recording will not be seem by anyone other than the primary researcher.

Any discomfort or stress beyond that standard to playing an instrument should be reported immediately.

BENEFITS OF THE STUDY

This study will inform future best practices in the field of music.

CONFIDENTIALITY OF RESEARCH INFORMATION

All of the information you provide will be confidential. However, if we learn that you intend to harm yourself or others, we must report that to the authorities.

OTHER INFORMATION

You may refuse to participate and you are free to withdraw from this study at any time without penalty or loss of benefits to which you are otherwise entitled.

No monetary compensation for participation.

RESEARCH-RELATED INJURY

If you think you have been harmed from being in this research, contact Doug Morin, Dr. Morrison, or Professor Salzman immediately.

The UW does not normally provide compensation for harm except through its discretionary program for medical injury. However, the law may allow you to seek other compensation if the harm is the fault of the researchers. You do not waive any right to seek payment by signing this consent form.

Subject's statement

This study has been explained to me. I volunteer to take part in this research. I have had a chance to ask questions. If I have questions later about the research, or if I have been harmed by participating in this study, I can contact one of the researchers listed on the first page of this consent form. If I have questions about my rights as a research subject, I can call the Human Subjects Division at (206) 543-0098 or call collect at (206) 221-5940. I will receive a copy of this consent form.

Printed name of subject   Signature of subject   Date

Copies to: Researcher

Subject

Appendix B: Initial Recruitment Email Sample

Dear Members of (ensemble name),

My name is Doug Morin, a Doctor of Musical Arts Candidate at the University of Washington. I have been given permission to send you this email by (ensemble director name).

I am asking for your participation in a study. This investigation looks into a brass or woodwind players' visual fixation patterns while they read music when a conductor directs. Participants must be an undergraduate Brass or Woodwind player, at least 18 years of age, and accepted into an auditioned concert ensemble.

Participants will receive a Starbucks gift card upon completion of the study.

Times to participate in the study will be set on an individual biases on (date) between 2 pm and 7 pm. The overall time commitment is approximately 20 minutes.

If you are interested, please click here to sign up for the study or paste this url into a browser – (insert url).

If you have further questions, please email me.

Sincerely,
Doug Morin

Disclaimer:
This study has no connection to any course at any university. Any decision to participate in this study is completely voluntary. The directors, instructors, and/or professors of record will not be notified of any decisions on whether or not an individual chose to participate. No personal information will be attached to any study results.

Appendix C: Example Solo