음악박사 학위논문

Cognitive Strategies of Sight-Reading in Piano Performance: Adapting the Eye–Hand Span

피아노 연주자의 초견 인지 전략: 눈–손 간격(EHS)을 중심으로

2024년 2월

서울대학교 대학원
음악과 음악학 전공
임 여 은

Cognitive Strategies of Sight-Reading in Piano Performance: Adapting the Eye–Hand Span

지도교수 오 희 숙

이 논문을 음악박사 학위논문으로 제출함

2024년 1월

서울대학교 대학원
음악과 음악학 전공

임 여 은

임여은의 음악박사 학위논문을 인준함

2024년 1월

위 원 장 ________이 교 구________(인)
부위원장 ________이 경 면________(인)
위　　원 ________오 희 숙________(인)
위　　원 ________박 정 미________(인)
위　　원 ________허 영 한________(인)

Abstract

This dissertation explores individual differences in sight-reading abilities among professional performers, with a particular focus on the eye–hand span. Sight-reading, which refers to reading and performing music for the first time without the advantage of prior rehearsal, is both a representative ability of a performer's outstanding musical talents and a fundamental skill that all musicians should acquire. Despite excellent instrumental techniques, there are substantial variations in sight-reading abilities among professional performers. This dissertation focuses on this phenomenon and addresses the following research question: Why do certain performers excel in sight-reading, while others face difficulties?

To explore this question, this dissertation measured the eye–hand span of professional pianists and examined its correlation with sight-reading proficiency evaluated in terms of performance accuracy. Eye–hand span, the distance between a performer's fixation and execution of a note, has long been considered a decisive indicator of sight-reading proficiency in the field of music psychology. However, the literature lacks an integrated perspective considering musical variables in the relationship between eye–hand span and sight-reading proficiency. Thus, this dissertation established three domains of sight-reading and investigated their interrelations. The domain indicators included musical complexity and playing tempo (musical domain), eye–hand span (cognitive domain), and performance accuracy (behavioral domain).

In the experiment, thirty-one professional pianists sight-read four musical pieces with two different complexities and playing tempi. This dissertation measured the participants' eye–hand span, evaluated their performance accuracy, and divided the participants into three groups based on their performance accuracy values. It investigated the variations in eye–hand span within each group and the influence of the musical domain on these variations. Surprisingly, the results showed that the eye–hand span did not change exclusively based on performance accuracy. Instead, the relationship between the eye–hand span and performance accuracy changed according to the difficulty of the sight-reading task, which is determined by both the complexity of the music and playing tempo. Notably, higher performance accuracy resulted in greater adaptability to this flexibility. The findings show that proficient sight-readers do not necessarily maintain a longer eye–hand span than do their less-skilled counterparts, but instead adjust their eye–hand span flexibly in response to moderating factors such as the characteristics of the music. Taken together, this dissertation reveals a shift in the framework of the eye–hand span from being a decisive indicator of sight-reading proficiency to understanding it as a cognitive strategy. This dissertation is significant as it demonstrates that the eye–hand span is not merely proportional to sight-reading proficiency but is rather a flexible strategy modulated by the musical context.

The scientific exploration of sight-reading offers insights on multiple fronts. From a cognitive science perspective, empirical studies on sight-reading enrich our understanding of multisensory integration and processing and the coordination of higher-order cognitive functions such as attention, memory, action, and prediction. Empirical research on proficient sight-reading also offers a window into how humans' expertise and skilled behavior are acquired, developed, and internalized.

From a pedagogical perspective, investigating the cognitive mechanisms of sight-reading can help educators devise evidence-based teaching methods and educational environments and develop personalized sight-reading training programs that cater to individual challenges of learners. From a musical standpoint, a systematic inquiry into sight-reading elucidates the specific skills that constitute superior sight-reading and contributes to the strategic refinement of sight-reading skills, which remains challenging among many professional performers.

Keywords: Music Cognition, Music Performance, Sight-Reading, Pianist, Eye Tracking, Eye–Hand Span (EHS), Complexity, Playing Tempo

Student Number: 2018-34350

Table of Contents

Chapter 1. Introduction 1

1.1. Definition and Delimitation 1
1.2. Motivation 4
1.3. Statement of Purpose 9
1.4. Outline of the Dissertation 10

Chapter 2. Background 12

2.1. Sight-Reading in Music Psychology 13
  2.1.1. Research History and Significance 13
  2.1.2. Core Processes and Components of Sight-Reading 16
  2.1.3. Factors Related to Sight-Reading Proficiency 25
  2.1.4. Improvement in Sight-Reading Skills 38
2.2. Eye Tracking in Sight-Reading 47
  2.2.1. Need of and Insights on Eye Tracking 47
  2.2.2. Findings of Eye-Tracking Research on Sight-Reading 51
  2.2.3. Eye–Hand Span 64
2.3. Conceptual Framework for the Study: The Three Domains 76
  2.3.1. Relationship Between Cognitive and Behavioral Domains 77
  2.3.2. Relationship Between Musical and Cognitive Domains 80
  2.3.3. Relationship Between Musical and Behavioral Domains 82
  2.3.4. Aims of the Study 84

Chapter 3. Methods 86

3.1. Participants 86
3.2. Sight-Reading Materials 87
  3.2.1. Standards of Musical Complexity 87
  3.2.2. Composition 87
  3.2.3. Measurement of Musical Complexity 94
3.3. Equipment 96
3.4. Procedures 96
3.5. Data Analysis 97
  3.5.1. Eye–Hand Span 97
  3.5.2. Performance Accuracy 99

Chapter 4. Results..........102

4.1. Performance Accuracy Based on Musical Complexity and Playing Tempo..........102
4.2. EHS Based on Musical Complexity and Playing Tempo..........106
4.3. Correlations Between the EHS and Performance Accuracy..........108
4.4. Correlations Between the EHS and Performance Accuracy Based on Sight-Reading Task Difficulty in High- and Low-Accuracy Groups..........110

Chapter 5. Discussion..........114

5.1. Main Findings and Implications..........114
  5.1.1. Relationship Between Cognitive and Behavioral Domains..........114
  5.1.2. Relationship Between Musical and Cognitive Domains..........118
  5.1.3. Relationship Between Musical and Behavioral Domains..........121
5.2. Limitations and Future Directions..........123

Chapter 6. Conclusion..........129

Bibliography..........133

국문 초록..........165

Acknowledgments..........168

List of Tables

Table 1. Summary of the eye-tracking literature on sight-reading 52

Table 2. Summary of the EHS literature on sight-reading........................ 68

Table 3. Quantitative schema of sight-reading materials.......................... 89

Table 4. Integrated, pitch, and rhythmic accuracy based on musical complexity and playing tempo 104

Table 5. *F* and *P* values of integrated accuracy with different musical complexities and playing tempi 104

Table 6. EHS with different musical complexities and playing tempi ... 107

Table 7. *F* and *P* values of the EHS with different musical complexities and playing tempi.. 108

Table 8. Correlation coefficients between the EHS (beat, sec, and note) values and performance accuracy (integrated, pitch, and rhythmic accuracy) values: Pearson correlation coefficients for integrated accuracy (IA); Spearman correlation coefficients for pitch and rhythmic accuracy 110

Table 9. Correlations between the EHS and integrated accuracy in high- and low-accuracy groups 112

List of Figures

Figure 1. A conceptual framework for the study 77

Figure 2.1. Piece 1, simple level 90

Figure 2.2. Piece 2, simple level 91

Figure 2.3. Piece 1, complex level 92

Figure 2.4. Piece 2, complex level 93

Figure 3. Comparison of the degree of pitch-class distribution between sight-reading materials and compositions of Western classical music composers 95

Figure 4. Visualization of the EHS calculation represented as a beat index 99

Figure 5. (a) Integrated, pitch, and rhythmic accuracy for different sight-reading tasks (simple-slow, simple-fast, complex-slow, and complex-fast). (b) Scatter plot depicting the regression of correlation between performances of simple and complex pieces in slow and fast tempi in terms of integrated accuracy. (c) Scatter plot depicting the regression of correlation between pitch and rhythmic accuracy. (d) EHS (beat, sec, and note) values for different sight-reading tasks (simple-slow, simple-fast, complex-slow, and complex-fast). (e) Scatter plot depicting the regression of correlation between the EHS (beat, sec, and note) and integrated accuracy. (f) Scatter plot depicting the regression of correlation between the EHS (beat, sec, and note) and pitch accuracy and correlation between the EHS (beat, sec, and note) and rhythmic accuracy 105

Figure 6. Correlations between the EHS (beat, sec, and note) and integrated accuracy across varying difficulties of the sight-reading task for high- and low-accuracy groups. (a) Correlation between the EHS (beat) and integrated accuracy. (b) Correlation between the EHS (sec) and integrated accuracy. (c) Correlation between the EHS (note) and integrated accuracy 113

Chapter 1. Introduction

1.1. Definition and Delimitation

This dissertation focuses on sight-reading. To establish a precise definition of what sight-reading is in the realm of music, it is crucial to examine how this concept has been defined in the broader literature. According to the *Harvard Dictionary of Music*, sight-reading refers to “the performing of a piece of music on seeing it for the first time” (Randel, 2003, p. 1417). Wolf (1976), who first proposed a cognitive model of sight-reading proficiency, described it as “the ability to play music from a printed score or part for the first time without benefit of practice” (p. 143). Since Wolf’s (1976) definition, numerous researchers have employed similar descriptions (Cara, 2018, 2023; Eaton, 1978; Gromko, 2004; Kopiez & Lee, 2008; Lörch, 2021; Tsangari, 2010). According to Thompson and Lehmann (2004, p. 145), the most common understanding of sight-reading involves “the practice of playing a piece of music directly from the score on first encounter or after brief rehearsal.” Lehmann and Kopiez (2016, p. 547) provided a more elaborate perspective, characterizing sight-reading as “the vocal or instrumental execution of longer stretches of non- or under-rehearsed music at an acceptable pace and with adequate expression.” Zhukov and McPherson (2022, p. 192) emphasized that “the first step in learning to play any notated musical work is to read through the score for the first time–prima vista or music sight-reading.”

These definitions emphasize two important components inherent in the term *sight-reading*: 1) The absence of prior practice and 2) the act of performing music. Specifically, the term is understood as reading music notations without previewing it beforehand and includes the act of listening to the performed music, allowing for real-time adjustments during the performance. This suggests that during sight-reading, visual input, motor execution, and auditory feedback all work together. Within the broader context provided by these definitions, this dissertation defines sight-reading as the ability to read and perform music for the first time—without the advantage of prior rehearsal. In this definition, the ultimate goal of sight-reading is assumed to be the execution of musical passages with the appropriate and natural musical flow and expressivity (Lehmann & Kopiez, 2016; Thompson & Lehmann, 2004), as sight-reading is fundamentally a musical performance.

To provide a comprehensive analysis of sight-reading, clearly delimiting the scope of sight-reading studies this dissertation addresses and narrowing the research focus are also essential. In this context, the following considerations apply. First, research on reading-only or nonperformance music reading, commonly called *silent music reading*, is excluded from this dissertation, as the focus is exclusively on sight-reading that involves actual performance. Although some may equate silent music reading—often simply termed *music reading* in the broader field of music psychology—with sight-reading, they are fundamentally different musical activities. Sight-reading, which in the literature conventionally implies *performance*, entails the reading and performance of music notation. By contrast, silent music reading omits the acts of performing and the resultant listening to produced sound. This distinction, marked by the absence of motor components and physical sound, results in varying cognitive demands compared to those associated with sight-reading (Puurtinen, 2018; Silva & Castro, 2019). Ortmann (1934), in his study of silent music reading, distinguished between sight-reading and silent music reading by stating that his experimental data could not be employed as an index for sight-reading, owing to the lack of reproduction of written musical notations and the auditory aspects of music. He highlighted the distinct characteristics of these reading modalities: "The ability quickly and correctly to see the visual stimulus does not include the abilities to reproduce it. However, the ability to *reproduce* [emphasis added by Ortmann] it does include the ability to read it quickly and correctly" (Ortmann, 1934, p. 57). Consequently, this dissertation concentrates solely on sight-reading studies while acknowledging the unique task requirements and cognitive mechanisms that distinguish it from silent music reading.

Second, this dissertation prioritizes the exploration of piano sight-reading within the broader landscape of instrumental and vocal sight-reading studies. One reason for this focus is the complexity of piano sight-reading. Pianists, while engaging in sight-reading, are tasked with the simultaneous interpretation of music notations from double staves, involving horizontal and vertical reading. This distinguishes piano sight-reading from the practices of other instrumentalists and vocalists, who typically perform from single-staff notations and horizontal reading. The dual reading requirement inherent in piano sight-reading necessitates *additional cognitive considerations* (Cox, 2000), rendering it a skill distinct from other sight-reading domains (Aiba & Matsui, 2016). In line with this specificity, most of the extant sight-reading literature has predominantly centered on piano sight-reading, surpassing other categories of sight-reading performers (Rayner & Pollatsek, 1997). Thus, this dissertation primarily addresses the literature on piano sight-reading, with minimal exceptions reserved for instances in which the research provides exceptional insights that are particularly pertinent to the relevant discourse.

1.2. Motivation

Why are sight-reading abilities considered significant in the realm of music? What drives the need to research this skill? Throughout the history of Western classical music, sight-reading has consistently held a pivotal role in musicians' skill sets. According to Lehmann and McArthur (2002), in the nineteenth century, musicians who were both composers and performers, such as Felix Mendelssohn (1809–1847) and notably Clara Schumann (1819–1896), popularized the tradition of *repertoire*. This tradition involved meticulous rehearsal and often the memorization of compositions by other musicians. Before this tradition took hold, performing music at first sight was a prevalent form of musical concerts. Carl Czerny (1791–1857) and Franz Liszt (1811–1886) were also recorded as publicly engaging in performances by sight-reading in their early years (Lehmann & Ericsson, 1996). Above all, superior sight-reading abilities have traditionally been considered a remarkable musical talent, often arousing an aura or mystique among performers (Lehmann et al., 2007). This mystique is well exemplified by the prodigious feats of Wolfgang Amadeus Mozart (1756–1791) during his childhood. Daines Barrington, an English magistrate and scholar, conducted tests on the young Mozart in London. His observations, reported in the Royal Society in 1770, confirmed Mozart's extraordinary sight-reading abilities. Barrington likened Mozart's sight-reading skills to an eight-year-old reading Shakespeare with the emotion of an actor and effortlessly and simultaneously deciphering commentaries in multiple languages, including Greek, Hebrew, and Etruscan scripts (Solomon, 1995).

As reflected in historical records, sight-reading is a fundamental musical ability that every musician should acquire among diverse musical aptitudes (Fan et al., 2022; Kopiez & Lee, 2006; Townsend, 1991) and is part of advanced musicianship (Kim et al., 2021). Sloboda (1978, p. 4) even stressed the musical significance of sight-reading, stating that the "[sight-] reading facility is not simply a useful additional skill for a musician to have. It is, in a sense, necessary for *full membership of the musical community* [emphasis added]." Various surveys also support this notion. For example, in Lowder (1983), which surveyed college faculty members and in-service teachers to identify the most crucial piano skills, sight-reading ability placed second in terms of importance, with cadences determined as the foremost skill. Following closely were score-reading, harmonization, and accompanying (Lowder, 1983, as cited in Kostka, 2000). Lyke (1968) devised a rating scale comprising 20 essential keyboard musicianship items and distributed it to both general music teachers and class piano instructors. The outcome was unanimous, with both groups ranking sight-reading as either the first or second most important skill. General music teachers ranked sight-reading and accompanying as the second and third most critical skills, respectively (Lyke, 1968, as cited in Watkins & Hughes, 1986).

Sight-reading has several advantages for performers. It offers access to a wider musical repertoire, accelerates the learning process of complex musical pieces, and reduces the preparation time for final performances (Dib & Sturmey, 2011; Russell, 2019; Wristen, 2005). It enhances musical independence, the sense of musical style, technique, and finger dexterity (Eaton, 1978). For advanced musicians, sight-reading makes it easier to participate in ensemble music performances and supports the academic analysis of music (Zhukov et al., 2016).

Sight-reading continues to serve as a prominent assessment tool for evaluating musical competence in various auditions and examinations (Chun, 2022; Waters et al., 1998). For example, the Associated Board of the Royal Schools of Music, a prestigious organization overseeing music examinations in the UK, assesses performance, technical, notation, and listening skills as part of their practical graded music exams, with sight-reading being a significant test component for evaluating *all-round musicianship* (Thompson & Lehmann, 2004).

Among various situations regarding sight-reading, skills remain of paramount importance, particularly for professional pianists in roles such as accompanists, repetiteurs, and orchestral pianists (Zhukov et al., 2019) because of some distinctive challenges during sight-reading. Collaborating with soloists, who frequently seek accompaniment for concert performances, places pianists in situations in which rapid familiarization with a wide range of repertoires under limited rehearsal time is essential. While aiding in the coaching of soloists, pianists are pressed to master pieces swiftly and adeptly to adhere to stringent time constraints, such as deadlines (Aiba & Matsui, 2016). However, despite acknowledging its significance, undergraduate piano students often report shortcomings in their sight-reading abilities and limited training opportunities (Zhukov, 2014). Taking these contexts into account, sight-reading is an increasingly crucial musical skill, particularly for professional pianists.

Professional pianists, despite their advanced musicality and technical expertise, exhibit individual differences in their sight-reading proficiency. These differences challenge the conventional assumption that pianistic virtuosity necessarily translates into adept sight-reading. An intriguing case is that of Josef Hofmann (1876–1957), a concert pianist and former Director of the Curtis Institute of Music in the US. Despite his exceptional musical aptitude, Hofmann was famously known for his poor sight-reading skills. According to Wolf (1976), who heard about Hofmann's sight-reading directly from his grandmother—with whom he had performed—Hofmann himself openly acknowledged this peculiarity. He rarely used printed music in concerts, even while performing with others. Whether playing solo or accompanying, he depended entirely on his remarkable musical memory. He could even handle orchestral reductions, such as the Bruch violin concerto, without sheet music. His poor sight-reading skills were attributed to his *extraordinary ear* for music, as he primarily learned by listening rather than reading notation (Wolf, 1976, p. 164). By contrast, Daniel Barenboim (born 1942), a world-renowned pianist and conductor, is recognized not only for his virtuosity but also for his accomplished sight-reading abilities. In his autobiography, *A Life in Music* (Barenboim, 2002), he describes himself as "reasonably good at sight-reading" (p. 68). As a result, he was able to "become acquainted with a vast amount of repertoire" (p. 5) and often accompanied singers in Lieder at first sight.

There are numerous cases in which even music majors with sufficient instrumental techniques and musicality were unable to excel in sight-reading (Banton, 1995; Bean, 1938; Iorio et al., 2023; Kornicke, 1992; Lehmann & Ericsson, 1993, 1996; McPherson, 1994; Mishra, 2014a, 2014b; Sloboda, 1984; Tsangari, 2010; Wolfs et al., 2020; Zhukov et al., 2016). Wolfs et al. (2020) showed that participants with adequate technical skills and a deep understanding of sight-reading materials still encountered fluency issues while performing at first sight. Mishra (2014b) suggested the possibility of a saturation point for certain constructs, such as music aptitude or technical ability, particularly among college musicians, where the impact of music aptitude on sight-reading seemed less pronounced than that of younger or less experienced performers, where it played a greater role. Furthermore, McPherson (1994) pointed out that sight-reading and performing rehearsed music should be considered separate and distinct aspects of musical performance, as evidenced by the differing correlations between both skills among participants with varying performance grades. This perspective asserts that proficiency in one domain does not guarantee excellence in the other, thus shedding light on the variability of pianists' sight-reading skills. Lehmann and Ericsson (1993), who involved expert pianists specializing in solo performance and piano accompaniment, effectively revealed consistent individual differences in sight-reading proficiency among professionals. Their study demonstrated stable variances in performance when expert pianists were tasked with sight-reading while accompanying prerecorded solo pieces. An interesting outcome of their study was the observation that the ability to sight-read did not exhibit notable improvements in increasing general instrumental skills among expert pianists. Surprisingly, pianists with a specialization in solo performance, who were assumed to be slightly more proficient overall, displayed inferior sight-reading abilities compared to that of their counterparts specializing in accompanying. In conclusion, their research accentuated discernible individual differences in sight-reading proficiency among expert pianists, contingent on their professional specialization. Kornicke (1992) comparably suggested that sight-reading and performing [rehearsed] music may require separate development, questioning the direct relationship between the two activities. These instances all support the idea that the ability to sight-read can greatly differ among professional pianists beyond the attainment of advanced pianistic skills.

1.3. Statement of Purpose

Building on the insights and issues derived from the historical and practical importance of sight-reading and evident individual variations in this skill, this dissertation tackles the intricacies of sight-reading proficiency. The primary objective is to scrutinize reasons behind individual differences in sight-reading abilities, particularly among professional pianists. The key research question is: Why do certain pianists excel in sight-reading while others struggle?

To answer this question, this dissertation employs a cognitive psychological lens. Sight-reading is not only an essential musical ability for performers but also a fascinating subject for scientific inquiry as it is a highly complex task orchestrating multiple sensory modalities and higher-order cognitive functions such as perception, memory, action, and prediction, all under demanding time constraints. This dissertation identifies the cognitive underpinnings of sight-reading proficiency and the scientific mechanics of this complex musical ability.

Specifically, this dissertation conducts an experimental study focusing on the eye–hand span (EHS) as it has been a measure closely linked to sight-reading proficiency in the field of music psychology. Through such empirical research, this dissertation aims to elucidate individual differences in sight-reading abilities in professionals and offer practical insights that can benefit performers, educators, and learners who are keen on refining their sight-reading skills and further musical capability.

1.4. Outline of the Dissertation

The remainder of this dissertation is structured as follows. Chapter 2 provides a theoretical background and context that guides the experimental study. It comprises three sections. The first section investigates the cognitive mechanisms behind sight-reading proficiency and practical strategies to enhance sight-reading skills. The second section narrows the interest to eye tracking and explores visual processing and visuomotor coordination in sight-reading. With a particular focus on the EHS, the third section scrutinizes issues and identifies gaps in the current research on the domain. It introduces a conceptual framework for the experimental study, which includes three domains of sight-reading, and outlines the specific aims of the study.

Chapter 3 details the methodologies, procedures, and data analysis techniques employed in the study. It presents novel approaches to quantify musical complexity and assess performance accuracy.

Chapter 4 presents the experimental results, supported by rigorous statistical analyses that yield comprehensive insights into the interrelations among the three domains of sight-reading.

Chapter 5 discusses the study's primary findings and their broader implications according to the conceptual framework introduced in the previous chapter. It critically evaluates the study's limitations and suggests avenues for future research on the EHS.

Chapter 6 revisits the key discoveries made throughout the study, offering possible answers to the research question regarding individual differences in sight-reading proficiency. It outlines future research directions, moving beyond the EHS to explore broader aspects of sight-reading proficiency. Finally, this chapter emphasizes the importance of a scientific perspective in understanding sight-reading, highlighting its cognitive, educational, and musical significance.

Chapter 2. Background

This chapter establishes a comprehensive background and sets the stage for the experimental study. It comprises three main sections. The first section briefly examines the history and significance of sight-reading research in music psychology and explores the broader discourse on sight-reading proficiency, addressing the core processes and components of sight-reading, factors influencing sight-reading proficiency, and pedagogical strategies that improve sight-reading skills. The literature discussed includes empirical studies at the behavioral level. An in-depth examination of psychological research oriented toward sight-reading proficiency offers theoretical clarifications of the enigma enveloping this proficiency.

The second section concentrates on eye-tracking research on sight-reading. The foremost action during sight-reading is reading music notation, and eye tracking is an invaluable methodology, providing tangible evidence of the differences in sight-reading proficiency during this procedure. This section underscores the necessity and significance of the eye-tracking approach in understanding sight-reading proficiency and scrutinizes various metrics and findings from eye-tracking studies, including eye movements, pupil dilation, perceptual span, and their relation to musical variables. An in-depth look at the EHS concludes this section. Among the measures related to eye tracking in the context of sight-reading proficiency, the EHS is a fundamental component encompassing the entire sight-reading process from the eyes to the hands. Thus, the

EHS is central to the main study conducted in this dissertation. An overview of the concept and relevant research on the EHS is provided.

The third section serves as a direct background for the experimental study. It introduces a conceptual framework that dissects sight-reading into three interrelated domains: musical, cognitive, and behavioral. Within the lens of the presented framework, this section examines the various issues and limitations evident in the extant EHS literature and contextualizes the need for this experimental study. It lists out the research objectives of the study.

2.1. Sight-Reading in Music Psychology

2.1.1. Research History and Significance

Sight-reading, a challenge for many performers, has also been a challenging but captivating subject for scholarly inquiry. As the introduction shows, it is a multisensory and highly complex cognitive activity. It involves encoding and decoding visual cues from the score, orchestrating complicated finger movements, and auditory monitoring of and adjustments to concurrent and subsequent performance. Above all, performers must undertake these processes under the strict constraints of real time.

The attempt to scientifically elucidate these intricate natures of sight-reading has persisted for nearly a century. Historical traces of the scholarly discourse on this subject can be found as far back as the 1930s (Bean, 1938). While earlier investigations in the 1920s touched upon subjects that were somewhat analogous to sight-reading—such as an eye movement study of music reading by Jacobsen (1928) and a study of sight-singing by Hillbrand (1924)—they were excluded as they fall beyond the scope of this dissertation. In the decades that followed, researchers such as Homer Weaver delved deep into the perceptual facets of sight-reading (Weaver, 1943). John Sloboda, in the 1970s and 1980s, investigated musical variables associated with sight-reading and the distinctions between expert and novice sight-readers, leaving behind profound contributions to our knowledge of sight-reading proficiency (Sloboda, 1974, 1976a, 1976b, 1977, 1978, 1984). From the 1990s onward, advancements in technology, notably eye tracking, facilitated nuanced studies on musical and nonmusical elements impacting sight-reading proficiency (Furneaux & Land, 1999; Goolsby, 1994a, 1994b; Kinsler & Carpenter, 1995; Polanka, 1995; Rayner & Pollatsek, 1997; Truitt et al., 1997;). The 2000s witnessed an enriched psychological discourse on sight-reading that was unprecedented in terms of depth and breadth; sight-reading began to be featured as individual chapters in various music psychology books (Lehmann et al., 2007; Lehmann & Kopiez, 2016; Lehmann & McArthur, 2002; Thompson & Lehmann, 2004; Zhukov & McPherson, 2022), coupled with multiple endeavors to construct predictive models explaining superior sight-reading skills (Gromko, 2004; Hayward & Gromko, 2009; Kopiez et al., 2006; Kopiez & Lee, 2006, 2008). The field has also seen neuroscientific approaches to the underlying mechanisms of sight-reading, adding another layer of depth to our understanding of the topic (D'Anselmo et al., 2015; Delogu et al., 2019; Lu et al., 2021).

The extensive body of sight-reading research points toward two overarching themes. One revolves around the intrinsic nature of sight-reading ability (Zhukov & McPherson, 2022), which encompasses the investigation of the sight-reading process and factors related to sight-reading proficiency (for detailed research trends in sight-reading, see Mishra, 2014b; Wristen, 2005 and the book chapters mentioned above). The other trajectory concerns the enhancement of sight-reading proficiency, primarily embracing the formulation and assessment of educational strategies aimed at continuously refining these skills (Alexander & Henry, 2012; Betts & Cassidy, 2000; Grutzmacher, 1987; Kostka, 2000; Mishra, 2014a; Salzberg & Wang, 1989; Watkins & Hughes, 1986; Zhukov, 2014, 2017; Zhukov et al., 2016). Both of these paths are important in understanding the complexity of sight-reading proficiency, and the subsequent reviews elaborate on each path.

The examination of sight-reading proficiency from a psychological viewpoint holds significant implications. In the realms of music perception and cognition, sight-reading is fundamental to understanding "the nature of music cognition itself" (Sloboda, 2005, p. 23), as it involves translating written notation into finger movements and reorganizing perceived notations with an intrinsic knowledge of musical structures in long-term memory. Studying sight-reading illuminates a performer's deep-seated understanding of music, providing deeper insights into the workings of the musical mind and its capacity. Expanding to a cognitive science perspective, sight-reading research deepens our comprehension of multisensory integration in complex cognitive tasks. Sight-reading requires a performer to rapidly convert visual stimuli into precise physical actions involving multiple high-level cognitive abilities. This transformative nature of sight-reading contributes to our understanding of sensory-motor integration and human cognitive capabilities. Pedagogically, an empirical approach to understanding sight-reading contributes to music education. By dissecting the cognitive steps involved in sight-reading, educators can develop differentiated training strategies that manage cognitive load and maximize the training effect, focusing on individual component skills of sight-reading. Educators can devise more personalized pedagogical methods that cater to individual challenges of learners. For example, exercises in the rapid and precise identification and execution of various musical patterns may be more effective for those lacking pattern-recognition skills. Such targeted and adaptive sight-reading training would allow students to refine their sight-reading skills efficiently, thus equipping them to handle the demands of live performance with greater confidence and competence.

2.1.2. Core Processes and Components of Sight-Reading

Before exploring the core processes and components of sight-reading, some key attributes that are integral to sight-reading are worth mentioning. First, sight-reading can be described as an *open skill* requiring constant adaptation to a changing musical environment (Lehmann & McArthur, 2002; Thompson & Lehmann, 2004). Unlike *closed skills* that are executed in stable conditions, such as rehearsed music performances or playing from memory, sight-reading demands that performers adjust in real time to unfamiliar music. Sight-reading can be likened to a *transcription task* in which written notations are swiftly converted into kinetic actions, much like reading aloud or copy typing (Fine et al., 2006; Sloboda, 1984). In this context, it involves a time lag between perceiving and translating a symbol into action. This skill is crucial for performers, as it demands the immediate performance of unfamiliar written music without prior rehearsal. Furthermore, sight-reading is an inherently *online activity* (Lehmann et al., 2007). Performers must generate sequences of movements in direct response to a continuous stream of visual stimuli, with the speed of the stimulus presentation determined by the tempo and the density of musical events. Unlike skills such as typing, sight-reading does not permit interruptions without disrupting the musical flow (Lehmann & McArthur, 2002). Finally, another intriguing attribute of sight-reading is that it is a *zero-sum game* regarding time allocation (Puurtinen et al., 2023). This accentuates the critical role of time in sight-reading performance, as performers must efficiently distribute their limited time among various processing tasks, while recognizing that allocating additional time to one symbol inevitably reduces that time for processing others. This zero-sum game aspect underscores the time-sensitive nature of sight-reading, where musicians must make rapid decisions while navigating dynamic musical scores. In summary, sight-reading is characterized by its adaptability to a changing musical environment, involving real-time conversion of notations into kinetic actions, all while requiring efficient time management.

What core mechanisms and psychological concepts underlie these attributes? This dissertation classifies the fundamental processes involved in sight-reading into perception, memory, action, and feedback, by drawing on insights from the literature (Drai-Zerbib et al., 2012; Lehmann & Kopiez, 2016; Lehmann & McArthur, 2002; Sheridan et al., 2020; Thompson, 1985; Thompson & Lehmann, 2004), thus establishing a structured understanding of the complexities of this skill.

Perception

For sight-reading, perception involves a dynamic interaction between performers' attention and the visual information in the musical score, laying the groundwork for subsequent cognitive processes (Thompson & Lehmann, 2004). In the course of perception, performers use eye movements to collect and integrate data into meaningful units and employ anticipatory eye movements to synchronize visual input with motor execution (Lehmann et al., 2007; Lehmann & Kopiez, 2016). Further elucidation on eye movements, a critical element of music notation perception, will be provided in the subsequent section.

Perception comprises two components. The first component, termed low-level perception routines (Thompson & Lehmann, 2004) or bottom-up processing, involves performers automatically encoding and decoding the physical properties of musical symbols, including pitches and durations (Lehmann & McArthur, 2002). The second component is higher-level cognitive functioning (Thompson & Lehmann, 2004) or top-down processing and is more closely related to the subsequent memory process. Here, performers draw on their stored knowledge in long-term memory to formulate hypotheses on the music's structure and anticipate its continuation, predicting the information necessary to effectively process the musical score (Lehmann & McArthur, 2002). The intricate interplay between these two fundamental components of the perceptual process serves as the gateway to musical interpretation and execution during sight-reading.

Memory

After perceiving musical notations, performers store, process, and retrieve relevant musical information in their memory before planning and translating it into appropriate kinetic actions. The role of memory in sight-reading is of paramount importance, as its effective operation in reorganizing encoded music notation, interpreting the piece, and planning for future motor execution has been established as closely linked to one's proficiency in sight-reading (Kopiez & Lee, 2008; Lehmann & Kopiez, 2016; Wolf, 1976).

The memory process in sight-reading features several critical components: Chunking, pattern recognition, and working memory. Chunking, a cognitive mechanism that finds application across diverse domains of expertise, involves grouping individual elements into more extensive and manageable units (Sheridan et al., 2020). It also contributes to optimizing the storage of this information in short-term memory (Gobet, 1998). In sight-reading, performers use chunking to structure visual inputs, such as musical notes or phrases, into coherent units. These chunks draw from the performer's existing knowledge, fostering more efficient processing of the musical score (Maturi & Sheridan, 2020). Notably, the size and organization of these chunks can vary and are influenced by the performer's level of expertise (Chase & Simon, 1973; Simon & Chase, 1988), thus affecting their ability to navigate musical passages fluently. Chunking is "in essence, a *memory mechanism* [emphasis added] that links our perception to previously stored knowledge" (Lehmann et al., 2007, p. 112).

Pattern recognition is an integral part of the memory process in sight-reading. While chunking focuses on how individuals organize and group musical information for easier processing, pattern recognition is the ability to identify familiar musical structures and relationships in music quickly. Through extensive practice, performers become adept at recognizing recurring patterns, such as arpeggios, chord progressions, and rhythmic motifs. In sight-reading, pattern recognition is the process through which current sight-reading experience activates perceptual representations stored in long-term memory, subsequently facilitating their matching with present perceptual experiences (Snyder, 2000, p. 23). This skill allows for rapid comprehension of music notation, particularly when faced with time constraints such as sight-reading (Bean, 1938; Cox, 2000; Wolf, 1976). Research has shown a positive correlation between pattern recognition skills and sight-reading proficiency, reaffirming the significance of pattern recognition in this context (Fine et al., 2006; Waters et al., 1998).

Finally, working memory, a cognitive system that serves as a temporary storage and manipulation hub for the information necessary to perform complex cognitive tasks (Baddeley, 1992), plays a crucial role in sight-reading (Herrero & Carriedo, 2019; Meinz & Hambrick, 2010). Working memory aligns with the idea of a cognitive workbench on which task-relevant data are actively held and processed (Lehmann et al., 2007). For instance, performers must recognize musical patterns or phrases and generate performance plans while using the notation as cues, anticipating how the music unfolds (Lehmann & Ericsson, 1996). Experts in sight-reading appear to organize their musical knowledge into retrieval structures within long-term working memory (Chaffin & Imreh, 2002; Drai-Zerbib & Baccino, 2014; Williamon & Valentine, 2002). These structures enhance their ability to access information efficiently and are particularly important for expert-level sight-reading (Drai-Zerbib et al., 2012; Ericsson & Kintsch, 1995).

In summary, memory is a fundamental component of the sight-reading process, encompassing chunking, pattern recognition, and working memory. These aspects collectively enable performers to store, process, and retrieve musical information, bridging perception and action.

Action

Following the memory process, the subsequent critical phase in sight-reading is action. In this phase, performers rely on specific tactile and kinesthetic skills to efficiently execute the needed motor movements. These skills allow them to orient themselves on their instruments without constantly requiring visual monitoring (Lehmann & McArthur, 2002), which can influence the perceptual process during sight-reading. When performers direct their gaze at the keyboard and fingers, doing so can disrupt their visual connection to the sheet music, resulting in a discontinuity in the continuous encoding of the music (Lehmann & Ericsson, 1996). Proficient sight-readers tend to glance at the keyboard less frequently (Cara, 2023; Gilman & Underwood, 2003; Lannert & Ullman, 1945).

In sight-reading, performers do more than execute predetermined motor sequences; they engage in preplanning movements, organizing intricate actions hierarchically within cognitive representations or maps, as described by motor program theory (Lashley, 1951, as cited in Thompson & Lehmann, 2004). This cognitive planning involves strategically mapping out movements and interpreting musical notations in real time, ensuring a coordinated and accurate execution of the piece. In this context, the orchestration of motor sequences and timing precision are critical (Drake & Palmer, 2000). Unlike tasks such as typing, where keystroke timing carries less weight, sight-reading demands playing the right notes in the correct order, making timing a crucial factor for excellence in sight-reading and broader musical performance. The internal timekeeper model provides valuable insights into the fine motor control of sight-reading. According to this model, the motor system can function as a timekeeper, overseeing the temporal aspects of movement trajectories and generating necessary time intervals. Musical performance has two timekeeper levels: One is responsible for meter pacing, and the other is integrated into the movement trajectories of note production, computed through motor procedures vis-à-vis the meter (Shaffer, 1982, 1984; Shaffer et al., 1985).

In essence, the action phase of sight-reading requires the integration of motor skills, cognitive planning, and precise timing, all of which contribute to the performer's ability to navigate through unfamiliar musical scores and deliver seamless performance during sight-reading.

Feedback and Prediction

After the action process, which involves the rapid translation of visual cues into motor commands, the final phase of the sight-reading process is feedback and prediction. These encompass auditory feedback and problem-solving mechanisms that performers employ to monitor and fine-tune their performance in real time. This is particularly crucial in ensuring that sight-reading performances are completed and polished.

Auditory feedback, as highlighted by Wolf (1976), serves as a real-time "verification mechanism" (p. 154), allowing performers to confirm that the transfer of information from visual input to finger movements proceeds smoothly and accurately. It helps detect discrepancies between what is expected and what is actually heard, guiding performers in making immediate adjustments. American pianist Vladimir Sokoloff (1913–1997) eloquently expressed this connection: "Part of sight-reading is the lateral process—as you play and as you hear what you play, that automatically drives you on to the next combination of notes and harmonies. *You can't divorce yourself from the sound* [emphasis added]" (Wolf, 1976, p. 155). Banton (1995) emphasized the role of auditory feedback in monitoring performance during sight-reading, showing that it is not only used but also significantly contributes to performers' ability to adjust their playing effectively.

Prediction is a simultaneous and integral component of the sight-reading process (Fine et al., 2006; Zhukov et al., 2016). Sight-reading transcends the mere interpretation of visual symbols on sheet music; it is a dynamic musical undertaking molded by the expectations cultivated through cumulative musical experiences and knowledge (Mishra, 2014a). In this regard, problem-solving skills are paramount, as performers must proactively anticipate potential challenges while concurrently deciphering musical notations. This anticipation helps them promptly recognize errors and inconsistencies, facilitating swift corrections guided by contextual cues within the music, their reservoirs of musical knowledge, and ongoing expectations during sight-reading (McPherson, 1994; Lehmann & Kopiez, 2016). The phenomenon known as *proof-readers' error*, exemplified by the case documented by Boris Goldovsky in Wolf (1976), offers a compelling validation of the predictive nature of sight-reading. In this case, a technically competent yet less proficient sight-reader played a G natural instead of the expected G# within a C# major chord, initially attributing it to a misprint. Conversely, more proficient sight-readers effortlessly "*inhibited* the G natural *in favor of the larger unit* [emphasis added by Wolf] (the C# major chord)" (Wolf, 1976, p. 169). This example underscores the fundamental role of predictive abilities and top-down processing in sight-reading (Sloboda, 1978). Viewed through this comprehensive lens, sight-reading can be linked to an inferential or improvisational process, akin to what

Bean (1938) described as music guesswork. This ability to create informed musical passages and fill in missing information, thus helps in the successful navigation of complex music scores. The acquisition of music-theoretical knowledge, including concepts such as harmony and counterpoints, along with familiarity with diverse musical repertoires and styles, significantly reinforces the predictive aspect of sight-reading. Proficient sight-readers, equipped with a deeper understanding of these facets, assume the ability to make precise predictions, thus reducing the cognitive load they experience during performances (Waters et al., 1998).

In summary, feedback and prediction emerge as pivotal elements within the sight-reading process, enabling performers to refine their playing through auditory feedback and problem-solving, encompassing anticipatory or inferential processes derived from their musical knowledge. This phase facilitates real-time adjustments, ultimately contributing to the overall success of the sight-reading endeavor.

This subsection has laid the groundwork for comprehending the fundamental processes and elements of sight-reading. It has described the dynamic and adaptive character of sight-reading, highlighting its real-time requirements and the intricate interplay among perception, memory, action, feedback, and prediction. With this foundational knowledge in place, the following subsection explores the factors associated with sight-reading proficiency and offers insights into how various factors, intricately woven into each psychological dimension of sight-reading, play a role in a performer's mastery over this complex skill.

2.1.3. Factors Related to Sight-Reading Proficiency

This subsection discusses factors related to sight-reading proficiency and how they are differentiated between skilled and less-skilled sight-readers. The discussion begins by examining factors related to each core process and component of sight-reading. This subsection investigates factors beyond these facets that contribute to individual variations in sight-reading proficiency. Finally, this subsection introduces predictive models that concurrently consider the role and impact of various factors on sight-reading expertise.

Perception

In relation to the perceptual process in sight-reading, proficiency relies heavily on visual perception and processing abilities, particularly those related to attention and the encoding of music notation. These abilities are essential for efficiently managing information that undergoes higher-level cognitive processing in memory, thus significantly enhancing skilled sight-reading (Fan et al., 2022). To investigate how skilled and less-skilled sight-readers perceive and process notational symbols differently, the most effective approach is to directly measure and compare their visual processing behaviors using eye tracking. Variables such as fixation, saccade, regression, and perceptual span offer valuable insights into the relationship between visual processing and sight-reading competence, shedding light on how these connections manifest at varying levels of proficiency. For example, proficient sight-readers can capture a greater number of notes within a single fixation than can their less proficient counterparts, thus enabling their perceptual ranges to expand further (Gilman & Underwood, 2003; Goolsby, 1994b). This heightened perceptual advantage is an ongoing area of investigation, providing a pathway to uncover the underlying cognitive mechanisms contributing to the “apparent superiority of good sight-readers in perceiving musical text” (Sloboda, 1974, p. 4). A more detailed discussion of how eye movements furnish a tangible framework for assessing and comprehending the various sight-reading skills within the context of the perception process will be fully presented in the second section of this chapter, titled “Eye Tracking in Sight-Reading.”

Memory

In the memory process, the discussion encompasses several crucial abilities contributing to sight-reading proficiency, that are related to each constituent part of this process: Chunking, pattern recognition, and working memory. Chunking is closely linked to sight-reading competence. Proficient sight-readers exhibit the ability to perceive larger musical patterns within notation than do less proficient ones (Kornicke, 1992; Sloboda, 1978; Wristen, 2005). This capacity enhances the efficiency of short-term memory usage and retrieval of musical information (Moussard et al., 2012; Rabinovich et al., 2014). Drawing an analogy between reading music and language, Bean (1938, p. 75) emphasized the importance of chunking in sight-reading proficiency, stating that “the good music reader must grasp units of four, six, or even eight notes just as he reads so many letters grouped together as a word. Therefore, appropriately grouped notes form units that are in visual or auditory perception, or in meaning as related to their context, the equivalent of words in prose or poetry. They should then be read as such.” Waters et al. (1998) suggested that skilled sight-readers can process larger musical units more quickly than can less-skilled ones by exploring the size of musical units or chunks processed by musicians while performing tasks, such as comparing two patterns comprising single-line melodies. In a study of professional pianists, Salis (1977) underscored the supremacy of chunking ability over tonal memory in sight-reading. Penttinen and Huovinen (2011) shed light on this concept by demonstrating that novice sight-readers often focus on individual notes and intervallic skips rather than rhythmical units. This approach demands more visual attention and thus leads to more errors during performance. In summary, expert sight-readers have the ability to effectively organize larger chunks of musical information. Efficient chunking strategies ease the reconstruction process in memory by alleviating the overall cognitive load (Bor et al., 2003; MacGregor, 1987; Norris et al., 2020).

Pattern recognition has consistently emerged as a critical factor in superior sight-reading skills (Polanka, 1995; Sloboda, 1978; Thompson, 1987; Wolf, 1976). The tachistoscope study by Bean (1938) provided compelling evidence: Skilled sight-readers demonstrated the capacity to perceive groups of three, four, or more musical notes at a single glance, distinguishing them as *pattern* readers who relied on whole-pattern cues for recognition, whereas those struggling with sight-reading were categorized as *part readers*, who relied on isolated notes and guesswork. Intriguingly, even in highly intricate compositions, accomplished sight-readers consistently displayed the ability to accurately read large musical patterns at a glance. Waters et al. (1998) corroborated these findings, demonstrating that the immediate recall of rapidly presented chords in pattern-recognition tasks exhibited the strongest correlation with sight-reading proficiency. Cox (2000) established this link, confirming a robust correlation between scores in chordal pattern reading and sight-reading and emphasizing the need for swiftly and accurately recognizing musical patterns for successful sight-reading. Skilled sight-readers exhibit a stronger reliance on pattern recognition, particularly while interpreting tonal music (Kim et al., 2021). The advantage of superior pattern-recognition skills in sight-reading lies in its facilitation of instant fingering selection through the recognition of standard fingering patterns in notation, resulting in heightened consistency and accuracy in performance (Sloboda et al., 1998; Zhukov, 2014).

Working memory also holds substantial importance in shaping an individual's sight-reading proficiency. The aspects of working memory have been thoroughly investigated in numerous studies that have aimed to unravel the complex cognitive mechanisms that underlie expert sight-reading. Herrero and Carriedo (2019) demonstrated that efficient retrieval and transformation subprocesses within working memory contributed to fewer errors, improved rhythmic accuracy, and enhanced expressiveness in sight-reading. In their study, working memory also aided in maintaining tempo, unaffected by task difficulty. Research that has explored the link between working memory capacity and sight-reading proficiency has yielded valuable insights into the role of working memory capacity in this domain. Working memory capacity, featured by its role in temporarily storing and processing task-relevant information, is frequently assessed through complex span tasks, such as those involving solving equations, while recalling other pertinent items, such as words (Engle, 2002). Meinz and Hambrick (2010) revealed that while deliberate practice accounted for a substantial portion (45.1%) of the variance in sight-reading performance, working memory capacity emerged as an additional influential factor, contributing significantly (7.4%) beyond the effects of practice. Importantly, their findings did not indicate that deliberate practice mitigated the impact of working memory capacity on sight-reading proficiency, suggesting that working memory capacity plays a direct and complementary role in determining sight-reading skills, potentially by enabling pianists to effectively anticipate and navigate ahead in music scores. Arthur et al. (2021) reinforced the role of working memory capacity in sight-reading, demonstrating that expert sight-readers had notably superior working memory capacity, as measured by the number of digits (in this study, numbers) recalled compared to both nonexpert sight-readers and participants without formal musical training backgrounds.

To summarize, individuals who excel at sight-reading appear to have better abilities in chunking, pattern recognition, and working memory capacity—key components of the memory process. Excellence in these facets seems to lead to efficient information retrieval and reorganization, and a seamless transition into the action domain during sight-reading.

Action

Sight-reading differs from silent music reading in that it engages in physical actions of *performance*. Therefore, examining how various motor skills are related to the act of performing at first sight can provide valuable insights into the mechanics of proficient sight-reading.

Expert sight-readers exhibit sophisticated motor programming patterns, allowing them to proficiently interpret familiar notational sequences (Wristen et al., 2006). As expertise grows, performers develop rule-based motor responses in their fingers, enabling them to combine movements into adaptable patterns that accurately correspond to musical notation (Wristen, 2005). Another significant aspect of successful sight-reading is fingering flexibility. Sloboda et al. (1998) demonstrated that master and expert pianists display a high degree of fingering flexibility, a critical component of their sight-reading skill. These skilled pianists do not rigidly adhere to standard fingering patterns. Instead, they have the capacity to choose fingerings that suit the musical context. This adaptability empowers them to navigate challenging passages, even when standard fingerings are not the most efficient or ergonomic choice. Expert pianists often have a repertoire of preferred fingering solutions for note sequences commonly occurring in specific musical contexts. By employing nonstandard finger spans and fingerings, expert sight-readers optimize their performance, ensuring smooth transitions between notes while maintaining tempo and expressiveness. This adaptability in fingering choices significantly contributes to their efficiency and accuracy when sight-reading diverse musical compositions (Sloboda et al., 1998).

Psychomotor skills emerge as vital predictors of sight-reading achievement. Eaton (1978) found that keyboard psychomotor skills, measured as the speed of key identification independent of years of experience, were the single most crucial predictor of sight-reading skill. Kopiez and Lee (2008) identified psychomotor speed, represented by tasks such as speed trilling and wrist tapping, as a strong predictor of sight-reading achievement.

Visual monitoring plays a critical role. As questions were raised by Sloboda (1978) regarding “at what level good [sight-] readers monitor their performance” (Sloboda, 1978, p. 12), it is noteworthy to understand that individuals who excel in sight-reading may rely on kinesthetic feedback and motor commands to detect errors before hearing them. Banton (1995) investigated the role of visual and auditory feedback in piano sight-reading, aiming to uncover errors resulting from feedback removal and compare pianists' performances based on their sight-reading practice frequency, musical experience, and ability. The study shows that visual feedback is essential in guiding precise hand movements, as its absence significantly increased adjacent note errors, underscoring its significance. However, the degree of reliance on visual feedback varied based on pianists' familiarity with sight-reading situations. Sight-reading without auditory feedback performed similarly to normal sight-reading, outperforming scenarios without visual feedback. These findings have implications for pianists' practice techniques and emphasize the importance of developing unguided keyboard navigation skills to enhance sight-reading fluency. Whereas Banton (1995) underscored the essential role of visual feedback in sight-reading proficiency, Cara (2018) further explored how proficient sight-readers adapt their visual monitoring to improve their performance. The author focused on the differences in glances at the keyboard (GAK) profiles between skilled and less-skilled sight-readers and found that the former exhibited fewer GAK per bar, often requiring fewer than two GAK per bar. More interestingly, skilled sight-readers adjusted their visual monitoring strategies based on the structural features of the music, such as musical accents and perceptual organization. This adaptability was reflected in faster GAK execution and increased anticipation across sight-reading trials, signifying their reduced reliance on visual feedback and improved GAK performance. These findings illustrate how visual monitoring evolves during sight-reading, with precision, planning, and speed as interconnected components contributing to superior performance.

In summary, the synthesis of advanced motor programming patterns, fingering adaptability, finely tuned psychomotor skills, and attentive visual monitoring collectively determine sight-reading proficiency in the execution process. These action-phase-related factors synergize to propel individuals toward becoming expert sight-readers.

Feedback and Prediction

In the realm of skilled sight-reading, auditory feedback is assumed to be crucial in performance monitoring. A study investigating the importance of auditory and visual feedback in sight-reading (Banton, 1995) found that performers without auditory feedback played comparably to standard sight-readers and even outperformed those without visual feedback. This suggests that auditory feedback has a minimal impact on the precision of motor movements in sight-reading. Instead, its primary function is performance monitoring. Successful sight-readers often initially underestimate their performance, whereas overestimations are linked to a higher number of errors. Proficient sight-readers tend to utilize auditory feedback when deviations from the intended sound occur, allowing them to make necessary adjustments. By contrast, less proficient sight-readers may struggle to create a vivid mental representation of the performance before engaging with the music, relying heavily on auditory feedback for validation, especially when motor demands are low. With increasing workload, their focus shifts from the produced sound to immediate information processing and initiating movements. Consequently, their initial memory of the performance emphasizes expected success, leading to overestimation before playback (Banton, 1995).

Aside from the role of auditory feedback, prediction assumes a central position in sight-reading proficiency. Waters et al. (1998) elucidated the significance of predictive processes by examining harmonic priming effects inspired by Bharucha (1987). Their research delved into whether highly skilled sight-readers derive greater benefits from the musical context than do their less-skilled counterparts. Through a chord-based priming paradigm, their findings revealed that proficient sight-readers demonstrate evidence of harmonic priming, indicative of their development of predictive mechanisms, such as sensitivity to chord regularities. These predictive skills significantly ease the cognitive load during sight-reading. Building on the proof-reader's error phenomenon observed by Boris Godovsky, Sloboda (1976b) conducted a structured inquiry on how performers handle incongruent notation by intentionally introducing misprints. Seasoned pianists were tasked with sight-reading these compositions and interestingly, despite their awareness and conscious efforts to spot errors, they consistently played what they expected to find in the score, rather than what was actually written. Sloboda's findings imply that proficient sight-reading involves the ability to anticipate probable musical continuations within a given idiom, rooted in a deep understanding of harmonic and rhythmic principles. This predictive capacity does not solely depend on technical expertise but rather on performers' adeptness at swiftly rectifying errors and self-monitoring performance (Chitalkina et al., 2021; Sloboda, 1976b, 1978).

In summary, auditory feedback and prediction mechanisms, encompassing performance monitoring and the anticipation of musical nuances, coexist as crucial factors related to sight-reading proficiency, with each contributing uniquely to performers' varying expertise in this intricate domain.

Further Insights

Sight-reading proficiency is shaped by a variety of factors that encompass core processes and additional considerations that elucidate the variations in individual differences in this complex skill. These factors can be broadly classified into domain-specific and domain-general influences. Within the former, background and experience hold significant sway (Arthur et al., 2020). Factors such as the extent and frequency of the sight-reading practice are key contributors to skilled sight-reading (Kornicke, 1992). Private instruction and the number of years dedicated to musical training, particularly for wind performers, emerge as strong predictors of sight-reading proficiency (Townsend, 1991). Lehmann and Ericsson (1996) noted that experience in accompanying and the breadth of accompanying repertoire are better predictors of sight-reading than is simply accumulating hours of piano practice. The capacity to self-regulate and employ effective strategies, including the identification of key and time signatures, mental rehearsal of challenging passages, and active engagement in error monitoring, intricately contributes toward sight-reading competence (Kim et al., 2021; McPherson, 1994). Rhythm reading skills have consistently played a pivotal role in sight-reading across multiple studies (Boyle, 1970; Elliott, 1982; Miller, 1988; Salzberg & Wang, 1989). For example, Elliott (1982) discovered a significant positive correlation between overall sight-reading competence and skill in reading rhythm patterns. Proficiency in rhythm reading stood out as a top predictor of sight-reading scores, with a combination of rhythm reading skills and performance jury scores being the most effective indicators of sight-reading aptitude, especially among college music majors. These findings corroborate McPherson's (1994) work, which highlighted the prevalence of rhythmic errors in sight-reading.

As for the latter, cognitive factors such as mental speed (Kopiez & Lee, 2008) and general intelligence (Luce, 1965; Salis, 1977) have been explored vis-à-vis sight-reading proficiency. Spatial-temporal reasoning, which is assumed to share a representational framework with music, has been proposed as an influence on sight-reading proficiency (Hayward & Gromko, 2009). Gromko (2004) revealed that text reading comprehension and spatial-temporal reasoning together explained a substantial portion of the variance in sight-reading skills.

This subsection explored the key factors associated with each process and component of sight-reading to provide an understanding of their individual significance. However, the complexity of sight-reading necessitates a more holistic perspective considering how all relevant factors interconnect to interpret sight-reading proficiency. Kopiez and Lee (2006) initiated this exploration by examining the relationship between predictor variables and the complexity of sight-reading tasks. They categorized these predictors into three distinct groups: General cognitive (e.g., working memory capacity, general mental capacity), elementary cognitive (e.g., speed of information processing, psychomotor speed), and expertise-related skills (e.g., sight-reading expertise, inner hearing). Their investigation involved 52 piano major graduates and undergraduates and featured sight-reading tasks of varying complexity. Their findings showcased the dynamic nature of sight-reading predictors, with specific combinations of these predictor variables gaining or losing significance as task complexity increased. For relatively simple sight-reading pieces, general pianistic expertise sufficed for success. However, as task complexity increased, factors such as psychomotor speed (measured as trilling speed), speed of information processing (measured as number combination test), inner hearing, and sight-reading expertise became progressively more critical. Interestingly, when the complexity of sight-reading task reaches its highest level, sight-reading expertise still plays an important role, but psychomotor speed becomes the main determinant of sight-reading achievement. This study concluded that sight-reading skills result from a nuanced interplay between skills honed through practice and influenced by innate abilities, such as psychomotor speed. It underscored the brain’s adaptability in optimizing these skills based on the demands of sight-reading tasks.

Building on these findings, Kopiez and Lee (2008) further categorized the component skills into general cognitive (e.g., short-term music memory, short-term numerical memory, working memory), elementary cognitive (e.g., speed of information processing, simple reaction time, tapping speed, psychomotor speed), and practice-related skills (e.g., inner hearing, accumulated hours of solo practice, piano lessons, sight-reading expertise) to gain a more comprehensive understanding of their correlation with sight-reading proficiency. They found that a combination of trilling speed, sight-reading expertise acquired before the age of 15 years, speed of information processing, and inner hearing were the most potent predictors, collectively explaining a substantial portion (59.6%) of the variance in sight-reading skills. This study illuminated the specific factors within each category that significantly influenced exceptional sight-reading skills and sought to discern how these factors interrelated and contributed to expert sight-reading. This emphasized the need to consider variables both related and unrelated to practice while assessing sight-reading proficiency, shedding light on the complex interplay among innate and experiential factors underpinning this skill.

Mishra (2014b) deepened the understanding of the predictors of sight-reading skills through an extensive investigation of the relationship between various variables and sight-reading proficiency across a broad spectrum of studies. Examining 92 studies resulted in 597 variables, which were grouped into constructs such as music aptitude, technical ability, improvisational skills, and ear-training ability. Separate meta-analyses were conducted for each construct, indicating variable effects on sight-reading. The findings suggested that skills that can be improved with practice, such as improvisational skills, ear-training ability, technical ability, and knowledge of music, were closely correlated with sight-reading, whereas stable characteristics such as attitude and personality showed no significant correlations. Mishra (2014b) highlighted that sight-reading is more than a simple visuomotor decoding process. Rather, it is a musical skill that improves with a performer's musicality. This insight implies the importance of musical understanding and construction prediction in sight-reading, directing future research toward these critical areas.

In conclusion, factors influencing sight-reading proficiency are multifaceted and include domain-specific and general influences. They contribute to the intricate landscape of sight-reading skills, with each component playing a unique and interconnected role in the overall success of sight-reading. Whereas factors pertaining to visual perception, memory processes, motor skills, feedback mechanisms, and predictive abilities are individually critical, their synergy drives individuals toward becoming proficient sight-readers. As researchers investigate how these elements interact and contribute collaboratively to the development of sight-reading proficiency, a more comprehensive understanding of sight-reading may emerge, enhancing the capacity to teach, practice, and appreciate the art of this complex musical skill.

2.1.4. Improvement in Sight-Reading Skills

Improving sight-reading skills is essential for performers at all levels, whether beginners, amateurs, or professionals. This skill proves valuable in situations that demand performing new scores *prima vista*, such as ensemble playing, auditions, or competitions, and enables performers to swiftly interpret musical scores and refine their final performances more efficiently while aesthetically mastering a particular piece. Although investigating the mechanisms of sight-reading and the proficiency-related factors delineated thus far can be considered an inquiry into the nature of sight-reading skills—and thus unveils the shroud surrounding them—determining ways to acquire and develop sight-reading skills constitutes a practical endeavor that is capable of bestowing substantial benefits on musicians seeking to deepen their sight-reading skills. This is important in the field of sight-reading studies (Penttinen & Huovinen, 2011).

Cognitive Strategies

Before exploring the methods and interventions and their effectiveness, it is noteworthy to first comprehend how advanced pianists employ a range of cognitive strategies when they play at first sight. According to Kim et al. (2021), who investigated expert pianists' sight-reading strategies across varying tonal environments, four primary operators—attention, static analysis, informed intuition, and performer's analysis—were found to be most important. Advanced professional pianists played a composition comprising tonal, nontonal, and ambiguously tonal sections, with immediate post-task reporting providing a window into their strategic approaches during the sight-reading task. Four strategies, attention, static analysis, informed intuition, and performer's analysis, emerged in common with each performer using these selectively based on the section's complexity. Participants often verbalized their strategies in sections that demanded heightened cognitive effort, possibly indicating a preference for economizing mental resources in less challenging segments. Static analysis, a key strategy, involved participants applying their existing musical knowledge to decipher structural elements within the music. When tonality was evident, participants primarily focused on recognizing tonal and stylistic elements. However, in sections with ambiguous or nontonal characteristics, their attention shifted to microlevel details, including the repetition of rhythmic patterns and nonharmonic tones. Informed intuition, proposed by Rink (1990, 2002), was another prominent strategy, as participants relied on their musical expertise to sense mood, texture, and tonal shifts. This intuitive approach extended to predictive skills, such as the anticipation of harmonic developments, particularly in tonal sections. The stability provided by a recognized tonal center emerged as crucial for enabling these predictive abilities. The study revealed two levels of performer's analysis. The first involved performers faithfully executing expressive markings in the score, whereas the second delved deeper, with participants adding nuanced expression beyond the explicit notation. The interplay between analysis and intuition for expressive performance challenged studies that suggested that only intuition was applied during sight-reading (Bangert et al., 2009). In summary, the exploration of sight-reading strategies exerted by advanced performers offers a keen perspective on effective techniques and approaches for higher-level sight-reading and a foundational understanding of how advanced performers independently overcome challenges and adopt problem-solving methods when confronted with unfamiliar musical scores.

Practical Guidance

Various recommendations and practical guidance for improving sight-reading skills have been presented in numerous psychological and educational studies (Alexander & Henry, 2012; Grutzmacher, 1987; Kostka, 2000; Mishra, 2014b, 2016; Pike & Carter, 2010; Russell, 2019; Salzberg & Wang, 1989; Watkins & Hughes, 1986; Wöllner et al., 2003; Zhukov, 2014, 2017; Zhukov et al., 2016). Before introducing the relevant discourse, it is essential to clarify that the upcoming strategies are built on the premise that sight-reading is a skill that one can be trained to acquire. The enduring debate over whether sight-reading ability is innate or acquired has persisted over time (Cox, 2000; Lehmann & Ericsson, 1996; Meinz & Hambrick, 2010; Mishra, 2014b; Townsend, 1991; Zhukov, 2017). Whereas inherent factors such as working memory capacity may undoubtedly influence exceptional sight-reading abilities (Arthur et al., 2021; Herrero & Carriedo, 2019; Meinz & Hambrick, 2010), most scholars have emphasized that sight-reading can be "developed through practice" (Colwell, 1969, p. 69, as cited in Townsend, 1991; for a comprehensive demonstration of the practice effect, see Platz et al., 2014). Sloboda (1978) alluded to the idea, drawing from the radical perspective of Buck (1944): "If you are a slow reader, remember that anyone can read a piece at a bar a minute, and there is no other excuse than laziness for not acquiring speed. If the child struggling with the cat sentence were to lament that she would never be able to read rapidly *like a grown-up* [emphasis added], you know that she is talking nonsense, and that the speed at which you and I can read English is not due to cleverness or any special gift. And the same is true, with no qualification, about the reading of music. To confess that you are a bad reader is to confess laziness ...” (Buck, 1944, as cited in Sloboda, 1978, pp. 3–4).

From this perspective, it is crucial to recognize that “deliberate practice is the central mediating mechanism to acquire skills at all levels and in different domains of expertise” (Lehmann & Ericsson, 1996, p. 6; see also Ericsson, 1996; Ericsson & Charness, 1994). In the sight-reading context, actual performance experiences in situations identical to sight-reading (e.g., under time constraints and with unfamiliar pieces) can significantly contribute toward enhancing sight-reading proficiency. Seeking appropriate challenges in this process is paramount (Lehmann & Ericsson, 1996). Considering these fundamental principles, this subsection elaborates on practical methods specifically designed to enhance sight-reading proficiency. This section first introduces the original opinions of authors who have written about sight-reading. By studying recommendations that account for a holistic view of sight-reading, performers can discern the relative importance among numerous strategies and specify the most beneficial approach to improve sight-reading skills.

Lehmann and McArthur (2002) addressed specific challenges such as misjudging intervals and provided remedies, including verbalizing interval names and scales, isolating interval recognition through flash-cards or drills, and focusing on melodic patterns before adding rhythm. To improve rhythmic skills, they recommended activities such as tapping or clapping rhythms, using metronomes, and practicing with live musicians. They recommended notating expressive elements from performances to enhance sensitivity to articulation and dynamics. To overcome common issues such as stuttering, they proposed exercises that encourage continuous play and discourage backtracking. Finally, they emphasized the importance of maintaining visual contact with sheet music during sight-reading.

Thompson and Lehmann (2004) presented sight-reading strategies by emphasizing the principles of deliberate practice. They advised developing a deep understanding of the musical style one wishes to play by sight-reading and suggested scanning new pieces for essential elements such as time and key signatures, phrase structures, and sections that may present difficulties. Inspired by McPherson (1994), the authors encouraged performers to apply the rules of musical expression they learned from analyzing rehearsed performances. They also recommended thinking critically about performance rules to refine sight-reading interpretations. Beyond the initial stages of improvement, they proposed identifying recurring problems and addressing specific weaknesses for more advanced sight-reading. They suggested recording and listening to one's performance, thus categorizing and analyzing errors to understand the nature of mistakes and determine whether a particular error pattern appears. Their perspective highlights the fact that talent has a limited role in sight-reading proficiency, with dedicated and inventive practice being the key to improvement.

Lehmann et al. (2007) emphasized the intriguing interconnectedness among sight-reading, memorization, and playing-by-ear skills among performers and that proficiency in one area can positively affect that in others. To bolster sight-reading skills, they proposed a proactive approach that includes participating in complementary activities and expanding one's repertoire with a deliberate focus.

While performers are encouraged to seek out challenging pieces and progressively tackle more complex music for the purpose of accompanying, mere repetition and more of the same may not necessarily lead to improvement. This approach significantly contributes to enhancing superior sight-reading skills, as it not only exposes individuals to diverse musical contexts but also cultivates their adaptability and problem-solving skills.

As a means to improve sight-reading skills, Lehmann and Kopiez (2016) recommended that young performers become accustomed to playing their instruments without constantly looking at their hands, which allowed them to thoroughly examine the score. Performing under real-time conditions without stopping for every mistake but rather inferring plausible content is also encouraged. The value of the theoretical knowledge of music and applying straightforward rules of expression to create musically sounding first impressions during sight-reading is emphasized.

Zhukov and McPherson (2022) suggested general and specific strategies for performers to ameliorate their sight-reading skills, emphasizing the need for purposeful training beyond mere regular practice. General strategies encompass the importance of consistent and systematic sight-reading practice, the benefits of duet playing to foster continuity and skill development, the selection of appropriate materials matching one’s sight-reading level, prioritizing pre-sight-reading preparations over immediate corrections to cultivate a positive attitude, and encouraging constructive self-evaluation and problem-solving. They stressed concluding sight-reading sessions on a positive note to maintain a long-term optimistic outlook. For specific strategies, experienced sight-readers can employ pre-sight-reading cues, such as scanning the entire piece to identify structural elements, comprehending key and time signatures, focusing on melodic shapes, vocalizing the melody, and interpreting tempo and dynamic markings for expressive playing. The authors advocated for clarifying uncertainties in the score before commencing and fostering knowledge transfer by comparing varying composers from the same period or style to promote pattern-recognition and prediction skills.

Integrated Training Approaches

As is evident in the aforementioned suggestions, research has underscored the importance of an integrated approach to enhancing sight-reading proficiency. Zhukov (2017) challenged the notion that experiential practice alone significantly improves sight-reading, highlighting the need for a multifaceted training approach. While classical university-level pianists participated in various forms of experiential sight-reading practice over a 10-week period, improvements in sight-reading performance were inconsistent across different performance indicators. Notably, aspects such as extra and missing notes, and beat adjustment showed minimal correlation with the overall duration of experiential practice. However, there was a silver lining. The study revealed that cumulative practice had a positive effect on one critical aspect, namely RMS accuracy, which measures the timing accuracy for each note played correctly. Therefore, Zhukov (2017) suggested that while experiential practice alone may not be the panacea for comprehensive sight-reading skills, there is potential for targeted improvements in specific areas. This emphasizes the need for a multifarious approach to sight-reading training, including training in rhythm and pitch, the cultivation of pattern recognition and predictive abilities, and engagement in collaborative musical activities.

Zhukov et al. (2016) proposed an innovative hybrid curriculum aimed at enhancing the sight-reading skills of advanced pianists. This curriculum combined three established teaching strategies, which were evaluated for their efficacy. The study involved 100 participants, organized into 4 groups of 25 each, that is, 1 control and 3 training (accompanying, rhythm, and style) groups. The results revealed significant advancements across all four sight-reading skill categories assessed (extra and missing notes, beat adjustment, and RMS accuracy) within the hybrid program, surpassing the progress observed in the individual training programs. By integrating rhythm training, exposure to diverse musical styles, and collaborative playing, this hybrid approach underscored that combining these strategies can be more beneficial than single-focus methods, providing valuable insights for both advanced pianists and educators.

Improving sight-reading skills can involve alternative approaches. Mishra (2016) emphasized the significance of customizing interventions to suit specific skill areas, incorporating counting systems, and integrating collaboration activities into sight-reading training. The study explored the impact of various interventions on rhythmic and melodic performance, revealing that these interventions had a small overall effect size, indicating modest enhancements in sight-reading skills. As for rhythmic sight-reading, treatments that centered on counting systems and included movement or rhythmic drills proved effective. In the context of melodic sight-reading, collaboration activities and instrumental training showed potential, although further research in this domain is necessary. Changes in notation did not improve sight-reading in rhythmic or melodic categories. These results suggest that interventions for sight-reading should take into account the distinct components of rhythmic and melodic sight-reading and adapt their approaches accordingly.

Individual Training Techniques

Finally, diverse individual studies enriched the repertoire of strategies available for improving sight-reading skills. Some focused on rhythm training, as observed in Zhukov (2014), whereas others explored aural development through exercises such as composition, keyboard harmony, and playing by ear, as advocated by Chun (2022). Aural imagery training that can be developed through early and regular sight-reading experiences, as discussed by Kornicke (1992), plays a significant role in achieving sight-reading proficiency. Certain studies addressed error detection and correction, as demonstrated by the emphasis in Kostka (2000) on error-detection practices and the use of prompts for rhythmic sight-reading, as examined by Salzberg and Wang (1989). Pattern recognition is another area of interest, with Grutzmacher (1987) investigating tonal pattern training and Pomerleau-Turcotte et al. (2023) highlighting the importance of acquiring knowledge on common harmonic patterns. Intriguingly, Watkins and Hughes (1986) demonstrated the positive effect of accompanying tape-recorded soloists on rhythmic accuracy during sight-reading.

In summary, improving sight-reading skills is essential for performers of all levels, extending advantages in the broader context of musical development and in specific contexts. An examination of advanced pianists’ cognitive strategies provides valuable insights into effective techniques and problem-solving methods toward superior sight-reading. The diverse methods and interventions presented in this section accentuate the teachable nature of this skill, addressing specific challenges and deliberate practice. Research has stressed the value of an integrated approach, underlining multifaceted training and collaborative activities. This wealth of practical strategies has equipped musicians with tangible means of enhancing their sight-reading skills and has strengthened their musical competence, enabling them to approach musical scores with increased confidence and artistry. This comprehensive discourse on improvement in sight-reading skills contributes significantly to performers' musicianship and their sight-reading practice, providing concrete strategies for success.

2.2. Eye Tracking in Sight-Reading

2.2.1. Need of and Insights on Eye Tracking

Thus far, this chapter has explored psychological discourses on sight-reading, including the fundamental mechanisms of it, factors related to sight-reading proficiency, and strategies for improving sight-reading skills. However, a crucial area that is yet to be addressed is eye tracking.

Eye tracking is an invaluable method for studying sight-reading, because *reading* musical scores is a foundational activity in sight-reading, and eye tracking serves as an optimal tool for examining human *reading* behavior. Eye tracking offers continuous, real-time recordings of reading performance, enabling analysis at global and local levels. Eye movements are inherent in the reading process, requiring no additional tasks from the reader. Eye tracking allows for an analysis of various measures of visual processing, including fixations, saccades, regressions, perceptual span, and the EHS. These measures closely mirror the processing demands associated with the structural features of written symbols and can capture individual differences among readers, such as reading proficiency (Raney et al., 2014; Rayner, 1997; Rayner et al., 2006, 2013).

The integration of eye tracking into sight-reading research is motivated by the extensive body of research in text reading, especially within the language domain. Several parallels between the reading of written language and music underscore the value of this integration. Structural and visual similarities between both domains (Sloboda, 1974, 1976a, 1976b, 1977), such as hierarchical organization (Cara & Gómez, 2016; Clarke, 1987; Koelsch et al., 2013) and left-to-right reading orientation (Hayward & Gromko, 2009), provide a compelling rationale. This shared cognitive ground extends to the processing of syntactic structures, as evidenced by comparable eye movement patterns (Ahken et al., 2012). The significance of grouping elements into larger units, such as phrases in music and language, enhances comprehension in both domains (Bean, 1938; Sloboda, 1977; Weaver, 1943).

The foundation of eye tracking in sight-reading research is built on the pioneering work of Jacobsen (1928) and Weaver (1943). They investigated eye movements vis-à-vis skill levels and types of music during sight-reading, respectively. Researchers have sought to explore the basic mechanisms of the oculomotor system and the unique cognitive processes involved in sight-reading (Chitalkina et al., 2021; Drai-Zerbib et al., 2012; Hadley et al., 2018; Huovinen et al., 2018; Kinsler & Carpenter, 1995; Puurtinen et al., 2023; Rayner & Pollatsek, 1997; Weaver, 1943) and have identified the optimal visual strategies employed by proficient sight-readers (Cara, 2023; Furneaux & Land, 1999; Gilman & Underwood, 2003; Goolsby, 1994a, 1994b; Penttinen et al., 2015; Polanka, 1995;

Qi & Adachi, 2022; Rosemann et al., 2016; Sloboda, 1974, 1977; Truitt et al., 1997; Waters & Underwood, 1998; Wurtz et al., 2009; Young, 1971; for a comprehensive review, see also Perra et al., 2021, 2022).

Historically, eye-tracking research on text reading has concentrated on various topics, including the characteristics and regulation of eye movements, the integration of information across saccades, and the perceptual span (Rayner, 1998). The focus shifted toward investigating individual differences in reading skills and the impact of varying task demands, particularly in the context of language acquisition and development (Ashby et al., 2005; Kuperman & Van Dyke, 2011; Radach & Kennedy, 2004, 2013; Rayner, 1986). This evolving emphasis aligns seamlessly with the shared interest in comprehending variations in sight-reading proficiency among performers.

The application of eye tracking in sight-reading offers a wealth of insights into the domain of music cognition and the broader context of cognitive research (Cara, 2018). From a music-cognitive perspective, eye-tracking data provide a window into performers' temporal and sequential processing while perceiving musical notations (Goolsby, 1994a), illuminating the cognitive mechanisms at play. For instance, how performers manage visual attention and prioritize elements within a score reflects their perceptual strategies, facilitates subsequent memory processes, and influences resultant performance outcomes. Studies have suggested that visual processing is closely linked to sight-reading fluency (Fan et al., 2022), with less proficient sight-readers potentially encountering challenges in tracking and decoding notational symbols over time. Eye tracking enables the investigation of the entire reading and performing music process. For instance, the EHS represents the distance between the eyes reading the sheet music and the hands playing the keyboard, which reveals a cognitive mechanism underpinning memory and the translation of memory into action during the sight-reading process. Through the EHS, investigating the efficiency of visuomotor coordination is possible—specifically, how collected visual information is stored, reconstructed, and converted into physical action. Eye tracking on sight-reading enhances the understanding of music cognition, as emphasized by Sloboda, who noted that "music [sight-] reading, despite its atypical input modality, is a *true species of music perception* [emphasis added]" (Sloboda, 1984, p. 224).

Beyond this music-cognitive perspective, applying eye tracking offers numerous insights into the broader cognitive context. It demonstrates how cognitive resources are allocated when humans perform complex tasks, especially those involving reading. Sight-reading is a prime example, requiring the simultaneous handling of multiple cognitive demands, including reading music. Managing cognitive resources in such a context deepens our understanding of task performance in music and across various domains demanding the coordination of multiple cognitive functions. Following this, eye-tracking studies in sight-reading have helped unravel the intricate interplay between visual expertise and the defining constraints and conditions, making it a valuable tool for understanding visual skills (Sheridan et al., 2020). The systematic and universally employed Western music notation provides a unique opportunity to explore the cognitive processes underlying symbolic systems, enriching our understanding of how the brain processes different types of symbolic information, such as text, music, and numbers, and the commonalities and differences in such processing (Madell & Hébert, 2008).

Building on these needs of and insights on eye tracking, the following subsections examine the findings of eye-tracking measures related to sight-reading proficiency and their implications. This examination is divided into two aspects: The first addresses general measures such as eye movements, including fixations, saccades, regressions, and other metrics such as pupil size and perceptual span. The second narrows the focus down further into the concept of the EHS, which serves as a specific measure of sight-reading proficiency.

2.2.2. Findings of Eye-Tracking Research on Sight-Reading

Table 1 summarizes the eye-tracking literature on sight-reading. Whereas studies that have measured the EHS are separately addressed in the subsequent subsection, "Eye–Hand Span," the focus here is solely on the literature that has investigated eye-tracking measures such as eye movements, pupil size, and perceptual span, excluding the EHS. The selection criteria for the review of eye-tracking research were as follows: 1) Studies measuring eye-related variables in a sight-reading situation (e.g., reading and playing music simultaneously), 2) Studies involving instrumental, at least vocal, sight-reading, and 3) Studies published in peer-reviewed and international journals. In all, 15 publications fulfilled these criteria.

Table 1. Summary of the eye-tracking literature on sight-reading.

Study	Participant			Skill Effect	Eye-Tracking Measure	Musical Variable	Sight-Reading Material			Playing Tempo Control
	Number	Type	Skill Level Classification				Type	Source	Length	
Goolsby (1994a)	24	12 skilled, 12 less-skilled sight-readers	Pre-task classification (sight-reading pretest)	O (*sight-singing)	Fixation, saccade, regression	Complexity	Single-line melody	4 melodies from *Solfege des Solfeges* by Danhauser, Lemoine & Lavigna (1910)	N/A	△ (tempo given prior to playing)
Truitt et al. (1997)	8	4 skilled, 4 less-skilled sight-readers	Post-task classification (playing time per measure)	O	Fixation, saccade, regression, perceptual span	Window size	Single-line melody	32 melodies from *Mikrokosmos*, Vol. 1 by Bartók (1940)	9, 12, 15, 18 bars	△ (tempo given prior to playing)
Gilman & Underwood (2003)	30	17 good, 13 poor sight-readers	Pre-task classification (sight-reading pretest)	O	Fixation, saccade, perceptual span	Window size	Double-stave music	32 phrases from Bach chorales	3 bars	X
Wurtz et al. (2009)	7		No level division	X (*violin sight-reading)	Fixation, regression	Complexity	Single-line melody	2 excerpts from sonatas by Corelli and Telemann	10 bars (Corelli), 20 bars (Telemann)	X

Study	Participant			Skill Effect	Eye-Tracking Measure	Musical Variable	Sight-Reading Material			Playing Tempo Control
	Number	Type	Skill Level Classification				Type	Source	Length	
Drai-Zerbib et al. (2012)	25	15 experts, 10 nonexperts	Pre-task classification (piano playing skill)	O	Fixation	Fingering difficulty	Double-stave music	36 piano excerpts from the classical tonal repertoire	4 bars	X
Penttinen et al. (2015)	38	14 performance majors, 24 education majors	Pre-task classification (musical background)	O	Fixation, gaze activity	Local structural feature (melodic alteration, rhythmic pattern)	Single-line melody	The children's song "Mary Had a Little Lamb"	8 bars	O
Arthur et al. (2016)	20	9 expert, 13 nonexpert sight-readers	Pre-task classification (piano pretest)	O	Fixation, saccade, regression	Visual disruption (spacing)	Single-line melody	10 newly composed melodies	4 bars	X
Huovinen et al. (2018)	37 (exp.1) 14 (exp.2)	14 professional, 23 amateur musicians (exp.1), 14 professional musicians (exp.2)	Pre-task classification (musical background)	O (only for exp.1)	Saccade	Local complexity (melodic skip)	Single-line melody	12 newly composed melodies (exp.1), 8 newly composed melodies (exp.2)	5 bars (exp.1), 24 bars (exp.2)	O

Study	Participant			Skill Effect	Eye-Tracking Measure	Musical Variable	Sight-Reading Material			Playing Tempo Control
	Number	Type	Skill Level Classification				Type	Source	Length	
Hadley et al. (2018)	24	Active pianists	No level division	X	Fixation, regression, pupil size	Anomaly (pitch)	Single-line melody	32 newly composed melodies	8 bars	X (exp.1) O (exp.2)
Cara (2018)	22	11 more skilled, 11 less-skilled pianists	Post-task classification (performance accuracy and speed)	O	Fixation	Complexity	Double-stave music	An excerpt from Ligeti's Etude No. 4, "Fanfares"	27 bars	X
Zhukov et al. (2019)	6	More experienced, less experienced players; better, weaker sight-readers	Pre-task classification (exam level of performance skill), Post-task classification (performance accuracy)	O (*woodwind sight-reading)	Fixation	Difficulty	Single-line melody	11 sight-reading examples from the *Watkins-Farnum Performance Scale* (1954)	N/A	O
Imai-Matsumura & Mutou (2021)	41	23 experts, 18 nonexperts	Pre-task classification (piano playing experience)	O	Fixation	Difficulty	Double-stave music	2 pieces from Sight Playing Workbook of YAMAHA Music Ability Test System	16 bars (easy), 30 bars (difficult)	X

Study	Participant			Skill Effect	Eye-Tracking Measure	Musical Variable	Sight-Reading Material			Playing Tempo Control
	Number	Type	Skill Level Classification				Type	Source	Length	
Lörch (2021)	144	74 music students, 70 hobby musicians	Pre-task classification (musical background)	O	Fixation, saccade, regression	Rhythmic pattern (note pair)	Single-line melody	48 newly composed melodies (4 sets of 12 melodies)	4 bars	O
Chitalkina et al. (2021)	22	Musically experienced performers	No level division	X (*both piano sight-reading and sight-singing)	Fixation, pupil size	Tonality, local congruence (melodic)	Single-line melody	"Mary Had a Little Lamb" by Lowell Mason	8 bars	O
Qi & Adachi (2022)	32	Advanced pianists	No level division	O (correlation analysis between performance error and eye-tracking measure)	Fixation	Modality (major, minor), complexity (intervallic)	Double-stave music	6 unknown pieces by Handel and Johann Ernst Bach	20 bars	X

Continuing from the criteria-driven selection as summarized in Table 1, two main approaches to the classification of participants' skill level consistently emerged across the eye-tracking literature on sight-reading: Pre-task and post-task methods. Through the pre-task approach, groups were delineated before the main experimental task began. Typical criteria included sight-reading pretests, as evidenced by Goolsby (1994a) and Gilman and Underwood (2003), or assessments of piano playing skills, such as those reported by Drai-Zerbib et al. (2012), Arthur et al. (2016), and Zhukov et al. (2019). Others took into account musical background and experience, such as education and academic position or degree, as seen in Penttinen et al. (2015), Huovinen et al. (2018), Imai-Matsumura and Mutou (2021), and Lörch (2021). By contrast, in the post-task methodology, the classification was conducted after the primary experiment. This classification often focused on parameters such as performance accuracy, as studied by Cara (2018) and Zhukov et al. (2019), or performance speed during the sight-reading task, as indicated by Truitt et al. (1997) and Cara (2018). An exception was Qi and Adachi (2022), which explored the relationship between sight-reading proficiency and eye movements, correlating performance errors with eye-tracking measures rather than adhering to a strict classification.

Some studies (Chitalkina et al., 2021; Hadley et al., 2018; Wurtz et al., 2009) only examined the influence of bottom-up factors (e.g., visual features of music notation, musical complexity, and task difficulty) on eye-tracking measures without considering the impact of top-down factors, such as sight-reading proficiency or performer's musical expertise. However, these studies were still included in the review. Whereas performers' sight-reading abilities or expertise—a top-down influence—undeniably affect visual processing, the inherent visual characteristics of the musical material play a significant role across all proficiency levels (Cara, 2023; Penttinen et al., 2013; Puurtinen, 2018). In the linguistic domain, visual complexities, encompassing factors such as word length or typographic density, directly influence saccade landing points and fixation durations, underscoring the integral relationship between perceptual anticipation and reading processes (Inhoff et al., 2000; Perra et al., 2021; Rayner et al., 2006; White, 2008). Concurrently, in the musical domain, readers demonstrate marked sensitivity to the visual elements of notations (Cara, 2018; Chitalkina et al., 2021; Drai-Zerbib et al., 2012; Goolsby, 1994a; Hadley et al., 2018; Huovinen et al., 2018; Penttinen et al., 2015; Puurtinen et al., 2023; Qi & Adachi, 2022; Wurtz et al., 2009). Distinctive features of musical notation, such as pitch elevations and rhythmic configurations, guide visual processing during sight-reading in a manner similar to the role of word familiarity in text reading (Penttinen, 2013). Such findings from eye-tracking research accentuate the preeminent influence of bottom-up factors on eye movement strategies during sight-reading (Zhukov et al., 2019). Given these insights, the review of the eye-tracking literature on sight-reading explores studies that have focused on top-down influences, such as sight-reading skills, and that have delved into bottom-up influences.

All studies have centered on piano sight-reading, except four: sight-singing by Goolsby (1994a); violin sight-reading by Wurtz et al. (2009); woodwind sight-reading by Zhukov et al. (2019); and both piano sight-reading and sight-singing by Chitalkina et al. (2021). Although the primary focus of this dissertation is on piano sight-reading, the other types are included in the review because of their implications and the limited scope of existing eye-tracking literature in the sight-reading field.

The eye-tracking measures investigated in the literature include eye movements, gaze activity, pupil size, and perceptual span. This review separately examines the results of these measures from the perspectives of top-down and bottom-up influences. First, considering skill level—a top-down-influence factor—eye movements emerge as the most frequently studied measure. Three major components of eye movements are prevalent during sight-reading: Fixations, saccades, and regressions.

Fixations are moments when the eyes remain relatively still, allowing the intake and processing of visual information. In general reading, individual fixations can vary in duration, ranging from a brief 50–100 ms to an extended 500 ms; however, they typically last approximately 200–250 ms (Rayner, 1977; Rayner & Pollatsek, 2006). These moments are critical for readers to encode and extract information. Most of the time, attention aligns with the point of fixation, providing precise clarity and focusing on a particular segment of the stimulus field (Rayner, 1998). On the other hand, saccades are swift eye movements connecting these fixations. While reading, saccades usually span 20–30 ms and extend over an average distance of 6–8 characters in English or approximately 2° of the visual angle (Rayner et al., 1981; Rayner & McConkie, 1976). As visual perception is momentarily inhibited during saccades, no new information is gathered from the text at this time (Rayner et al., 2006). Notably, not all saccades progress linearly during reading. Approximately 10–15% of the time, eyes revisit text sections that are already seen, a movement called regressions (Rayner, 1999). Whereas many regressions are brief—often directing the gaze to the immediately preceding word or retracting just a few letters owing to overextension or current word processing challenges—others can be considerably lengthier.

In general reading contexts, improved reading proficiency corresponds to specific eye movement patterns. For example, advanced readers typically exhibit fewer fixations, shorter fixation durations, longer saccades, and fewer regressions (Krieber et al., 2016; Rayner, 1986; Rayner et al., 2006; Reichle et al., 2013). By contrast, beginners and those with reading difficulties—including dyslexics—often demonstrate more fixations, longer fixation durations, shorter saccades, and increased regressions (Starr & Rayner, 2001). These patterns suggest that skilled readers efficiently encode words and use parafoveal and peripheral information, whereas less proficient ones face challenges in these areas.

As for sight-reading, similar trends are observed. Skilled sight-readers generally have fewer fixations (Gilman & Underwood, 2003; Lörch, 2021; Qi & Adachi, 2022; Zhukov et al., 2019) and shorter fixation durations (Drai-Zerbib et al., 2012; Goolsby, 1994a; Imai-Matsumura & Mutou, 2021; Lörch, 2021; Penttinen et al., 2015; Truitt et al., 1997; Zhukov et al., 2019). They also tend to have longer saccades (Huovinen et al., 2018) and fewer regressions (Drai-Zerbib et al., 2012). However, some studies have reported opposite patterns: More fixations (Goolsby, 1994a), longer fixation durations (Qi & Adachi, 2022), shorter saccades (Goolsby, 1994a), and more regressions (Goolsby, 1994a) for skilled sight-readers. Some have even indicated no difference between skilled and less-skilled sight-readers concerning the number of fixations (Arthur et al., 2016; Cara, 2018), fixation durations (Arthur et al., 2016; Cara, 2018; Gilman & Underwood, 2003), and saccades (saccade length in Gilman & Underwood, 2003; latency, speed, and number of saccades in Arthur et al., 2016). Nevertheless, a recent meta-analysis on eye movements across varying levels of expertise confirmed the robust trend of reduced fixation duration in experts, regardless of the type of music being sight-read (Perra et al., 2022). The development of sight-reading skills is indeed associated with the immediate recognition of note symbols reflected in decreased fixation time (Penttinen & Huovinen, 2011). These findings suggest that skilled sight-readers process musical information more rapidly and systematically than do their less-skilled counterparts, allowing them to recognize extensive musical patterns, reduce frequent pauses, and conserve mental resources for better performance (Lörch, 2021; Perra et al., 2022; Zhukov et al., 2019).

Another common aspect of eye-tracking research on sight-reading is perceptual span. Often termed the visual span or effective visual field, perceptual span represents the area where visual information is accessible and processed during a single fixation. This region extends beyond the fovea, capturing visually perceived symbols that appear blurred (Burman & Booth, 2009; Jacobs, 1986; Puurtinen et al., 2023; Rayner, 1998, 2009; Sheridan et al., 2020). Perceptual span serves as an index, providing insights into the amount of information entering the reading system at a specific moment (Madell & Hébert, 2008; Perra et al., 2021; Sheridan et al., 2020). A popular method to measure perceptual span is the moving-window paradigm or gaze-contingent window paradigm introduced by McConkie and Rayner (1975). In this approach, a text or stimulus is presented, and only a specific window around the point of fixation is fully visible, whereas the rest is masked or occluded. This window aligns with the reader's eye movements. By adjusting the window size, researchers can determine the minimum size needed for uninhibited reading, which then indicates the perceptual span (Gilman & Underwood, 2003; Perra et al., 2021; Rayner & Pollatsek, 1997; Truitt et al., 1997).

In language research, reading proficiency appears to impact the perceptual span. Skilled readers—characterized as college-age or faster readers, or those with superior reading comprehension and spelling abilities—exhibit a larger perceptual span than do less-skilled readers, such as beginning, older, slower, and dyslexic readers (Häikiö et al., 2009; Rayner et al., 1989, 2010; Veldre & Andrews, 2015). A broader span indicates an adept reader's capability to swiftly identify central words and subsequently allocate greater attention to nearby parafoveal content. Enhanced reading proficiency amplifies visual intake, equipping proficient readers to process and capitalize efficiently on information extending beyond their primary focus (Bélanger et al., 2012; Choi et al., 2015).

Interestingly, in the context of sight-reading, no significant difference in the size of the perceptual span between skilled and less-skilled sight-readers was observed (Gilman & Underwood, 2003; Truitt et al., 1997). Interpreting this phenomenon, Gilman and Underwood (2003) suggested that while skilled sight-readers may inherently have broader perceptual spans for musical material, they intentionally limit their spans during sight-reading because of the high demands of working memory in time-constrained tasks, where reading further into the periphery could prove detrimental.

As previously mentioned, visual processing during sight-reading is shaped by multiple factors. This mechanism can vary based on the individual—who is reading (a top-down influence)—and the material—what is being read (a bottom-up influence) (Puurtinen, 2018). Numerous studies have investigated how various bottom-up factors, coupled with the skill level of sight-readers, affect this process. The key musical variables examined include complexity (Cara, 2018; Goolsby, 1994a; Huovinen et al., 2018; Wurtz et al., 2009), difficulty (Imai-Matsumura & Mutou, 2021; Zhukov et al., 2019), tonality (Chitalkina et al., 2021), fingering (Drai-Zerbib et al., 2012), local musical features such as melodic incongruence

(Chitalkina et al., 2021) and alterations (Penttinen et al., 2015), and rhythmic patterns (Lörch, 2021; Penttinen et al., 2015), and others, including disrupted spacing (Arthur et al., 2016) and pitch anomalies (Hadley et al., 2018).

Given the diverse experimental methods and settings across studies, summarizing the effects of bottom-up influence is not straightforward. However, a recurring observation is evident: Music with higher complexity (Wurtz et al., 2009) or difficulty (Imai-Matsumura & Mutou, 2021) or unconventional and unexpected progressions (Chitalkina et al., 2021; Hadley et al., 2018; Lörch, 2021) usually results in longer fixation durations and increased pupil sizes. This observation aligns with the findings in language research, where text properties such as passage difficulty or inconsistency have been correlated with longer fixation durations and pupil dilation (Hess & Polt, 1964; Krejtz et al., 2018; Rayner et al., 2006). These results emphasize the varying cognitive loads and processing efficiency in sight-reading based on the music's structural or visual attributes.

Finally, Kinsler and Carpenter (1995) presented an eye movement model for sight-reading. Although their model was pioneering, it employed a simplified example and relied on non-instrumental tasks, such as rhythm tapping. The model comprises three phases: Encoding, processing, and executing. The encoding phase is the initial step in which the visual representation of musical notation is converted into neural activity. This phase does not delve into the interpretation of musical symbols. Rather, it is a preliminary transformation of visual patterns into neural signals. Following the encoding phase, the visual information is interpreted musically in the processing phase. The model posits that different notes may require varied interpretation durations based on their complexity and the proficiency of the performer. In this phase, the processor interprets the iconic representation of the fixated image, sets an accuracy criterion, and determines the next fixation point (Waters & Underwood, 1998). The concluding executive phase is where the interpreted musical data are translated into specific muscular commands needed for the actual performance of the piece. This phase can function autonomously, as music can be read and memorized without being performed; similarly, execution does not rely on visible notes. The duration of this phase can fluctuate based on the tempo and complexity of the musical piece (Kinsler & Carpenter, 1995). The buffer, akin to working memory, stands central in their model, acting as the bridge between the processor and executive unit (Madell & Hébert, 2008).

Nevertheless, Kinsler and Carpenter's (1995) model had limitations. The model was primarily designed for saccadic eye movements while reading basic note sequences, neglecting the nuances of reading and performing a full musical piece or the complexities of reading chords or multiple music lines at once. Huovinen et al. (2018) critiqued the model for its reliance and informal observations linked to basic rhythmic tasks, emphasizing that its narrow scope, given that it was based on only four participants, may exaggerate certain aspects. In line with this, Puurtinen et al. (2023) proposed three potential cognitive mechanisms involving visual processing during sight-reading. First, *symbol comprehension* suggests that the complexity of a musical symbol mainly affects a reader's foveal processing. Next, *visual anticipation* pertains to the proactive behavior of sight-readers, wherein they look ahead to prepare and plan their action. Finally, *symbol performance demands* are rooted in the multitasking aspect of sight-reading, positing that the intricacy of the symbols currently being performed can dictate how upcoming symbols are visually processed.

2.2.3. Eye–Hand Span

The eye-tracking measures discussed thus far provide valuable insights into the visual processing aspect of sight-reading. However, focusing solely on eye-tracking metrics without considering the motor aspect may not be sufficient to fully understand the complete transition from visual intake to performance output (Iorio et al., 2023). Sight-reading is an inherently multisensory activity that intricately integrates visual, auditory, and motor components. It also engages in a transcription task, converting notational information into kinesthetic action (Fine et al., 2006). The quality of the sound produced is a primary determinant of proficient sight-reading. As Wolf (1976) emphasized, the focus should be on the cognitive processing that converts the visual image into a muscular act rather than solely on what the performer visually perceives. For some individuals, the translation of visual patterns or meanings to the keyboard remains a challenging, unresolved matter (Bean, 1938).

A crucial aspect to consider is the EHS, which represents the distance between where a performer's eyes fixate on the score and the note they are playing. The EHS serves as a distinguished eye-tracking measure associated with individual differences in sight-reading abilities. When performers play at first sight, there is an interval between perception and action. This delay is integral because it facilitates the performer's processing of forthcoming notes and potential challenges, priming them for precise execution (Perra et al., 2021). As the performer's gaze typically anticipates the note to be played, this interval maintains a delicate balance. It promotes anticipatory movement planning while ensuring that the incoming visual data do not exceed the performer's memory limits, which helps prevent potential cognitive overload (Cara, 2018; Penttinen, 2013; Rayner & Pollatsek, 1997). Echoing the Kinsler and Carpenter (1995) model, performers synchronize extracted musical information with motor actions using a buffer system. The EHS illuminates how long visual information is retained in this buffer before being translated into finger actions, providing insights into the role of working memory in sight-reading proficiency (Furneaux & Land, 1999).

Given the significance of the EHS, this dissertation focuses on it in both the ensuing discussions and the subsequent experimental study. This approach aims to shed light on the real-time interplay between perception and action in sight-reading, which is crucial for understanding sight-reading proficiency.

The concept of the EHS in sight-reading is rooted in language and takes inspiration from the eye-voice span (EVS), which indicates the interval between visual tracking and spoken articulation during verbal reading (Inhoff et al., 2011). This notion can be traced back to Quantz (1897), who was the first to quantify this span. He used a method that involved observing the number of words a person read before a card blocked the text. He noted that this span generally expanded more at the beginning of a line than its end and could diminish to zero in circumstances such as while encountering an unfamiliar word (Weaver, 1943). Subsequent studies by Buswell (1920) and Butsch (1932) utilized photographic methods to determine the EVS and EHS in oral reading and typewriting, respectively. Levin and Kaplan (1970) introduced a technique to measure the EVS by extinguishing the light illuminating a text while participants read aloud. Afterward, the participants recounted the words they visually processed beyond the last word they spoke before the light was turned off. The EVS is characterized by the amount of content that individuals can accurately recall after the text becomes unavailable (Sloboda, 1977). Research on the EVS indicates that it is generally longer for skilled readers (Tinker, 1958) and for structured content, such as sentences, than it is for unstructured content, such as word lists (Lawson, 1961; Morton, 1964, as cited in Sloboda, 1974).

Drawing from the EVS, the measurement of EHS in sight-reading began with the pioneering study of Weaver (1943), which utilized a photographic approach. Early techniques often involved lights or photographic slides of sight-reading materials unexpectedly shutting off mid-trial (Sloboda, 1974, 1977). As eye-tracking methods evolved, measuring span size became more sophisticated (Furneaux & Land, 1999; Goolsby, 1994b; Truitt et al., 1997). Although there have been limited studies, consistent advancements have been observed in the field. Notably, as of 2023, interest in EHS has surged; out of the EHS papers published over nearly 80 years, over half have been written in the past decade.

Before exploring the extensive EHS literature, an important distinction must be made. While the conceptual foundation of EHS may be inspired by the methodologies of EVS, directly comparing EHS and EVS research findings can be misleading because of inherent differences between [music] sight-reading and language reading. In sight-reading, particularly in Western musical traditions, musical notations require performers to align strictly with temporal guidelines, ensuring synchronization with a collective tempo (Huovinen et al., 2018; Penttinen et al., 2015). Disrupting the rhythm can alter the message that music conveys. By contrast, text reading permits intermittent pauses, preserving the essence of the message (Silva & Castro, 2019). Sight-reading requires tracking multiple melodic lines at a consistent tempo, a complexity not seen in “superficially similar skills” such as reading aloud or typewriting (Thompson & Lehmann, 2004, p. 146; see also Bean, 1938; Weaver, 1943). The ability to read aloud generally does not present a significant challenge once a certain school age or literacy level is reached. By contrast, the ability to sight-read remains a considerable challenge for many *professional* performers, even those with extensive musical knowledge and advanced performance skills (Wolf, 1976). Therefore, individual variability in sight-reading skills among performers may be more pronounced than that in reading-aloud skills among speakers. Language research often prioritizes reading comprehension, focusing on visual processing aspects such as eye movements, more than visuomotor coordination represented by the EVS. However, the EHS, encompassing the entire perception-action process, offers deeper insights into reading and performing music at first sight. Considering the significant differences among tasks such as sight-reading, reading aloud, and typewriting, reviewing the findings of EVS in this dissertation may be deemed unnecessary. For these reasons, the subsequent review focuses exclusively on the EHS.

Table 2 provides a comprehensive summary of the EHS literature on sight-reading. As with the previous review of eye-tracking literature on this topic, the selection criteria include studies: 1) Measuring the EHS in the context of sight-reading, such as simultaneous reading and performing music, 2) encompassing instrumental sight-reading, and 3) published in peer-reviewed and international journals. In all, 16 publications met these criteria.

Table 2. Summary of the EHS literature on sight-reading.

Study	Eye Tracker	Participant			Skill Effect	EHS				Musical Variable	Sight-Reading Material			Playing Tempo Control
		N	Type	Skill Level Classification		Note	Time	Beat	ETS		Type	Source	Length	
Weaver (1943)	X (photographic method)	15	Well-trained musicians (9 of 15 were professionals)	No level division	X	1.9-3.1 (1.5 for chord)				Texture (harmonic, melodic, single-melody-with-supporting-chord)	Double-stave music	3 excerpts (excerpt 1 from a hymn, excerpt 2 from a minuet by Bach)	8 bars	X
Sloboda (1974)	X	10	Most music students	Subjects covered a wide range of sight-reading ability (but not specified)	O (correlation analysis between performance ratings and EHS)	3.8-6.8				Phrase boundary	Single-line melody	15 little known English and French popular folk melodies	N/A	O
Sloboda (1977)	X	6	Accomplished keyboard sight-readers	No level division	X	4.33-6.02				Phrase marker (physical, structural)	Single-line melody	A simple diatonic folk- or hymn-style melody and its variations	5, 7, 9 notes	X

Study	Eye Tracker	Participant			Skill Effect	EHS				Musical Variable	Sight-Reading Material			Playing Tempo Control
		N	Type	Skill Level Classification		Note	Time	Beat	ETS		Type	Source	Length	
Truitt et al. (1997)	O	8	4 skilled, 4 less-skilled sight-readers	Post-task classification (playing time per measure)	O			1-2		Window size	Single-line melody	32 melodies from *Mikrokosmos*, Vol. 1 by Bartók (1940)	9, 12, 15, 18 bars	△ (tempo given prior to playing)
Furneaux & Land (1999)	O	8	3 novices, 3 intermediates, 2 professionals	Pre-task classification (both exam level and self-report)	O	2-4	1 s			No variable	Double-stave music	5 excerpts from already published under a particular grade standard	N/A	△ (metronome being silenced after playing)
Gilman & Underwood (2003)	O	30	17 good, 13 poor sight-readers	Pre-task classification (sight-reading pretest)	O			0.75 -1		Window size	Double-stave music	32 phrases from Bach chorales	3 bars	X
Wurtz et al. (2009)	O	7		No level division	X (*violin sight-reading)	3-6	1 s			Complexity	Single-line melody	2 excerpts from sonatas by Corelli and Telemann	10 bars (Corelli), 20 bars (Telemann)	X

Study	Eye Tracker	Participant			Skill Effect	EHS				Musical Variable	Sight-Reading Material			Playing Tempo Control
		N	Type	Skill Level Classification		Note	Time	Beat	ETS		Type	Source	Length	
Penttinen et al. (2015)	O	38	14 performance, 24 education majors	Pre-task classification (musical background)	O		1 s	1-3		Local structural feature (melodic alteration, rhythmic pattern)	Single-line melody	The children's song "Mary Had a Little Lamb"	8 bars	O
Rosemann et al. (2016)	O	9	Piano major students	No level division	O (correlation analysis between performance ratings, musical experience, and EHS)		1-1.5 s	0.5		Complexity	Double-stave music	The accompaniment part from 'Adagio ma non tanto' of Bach's Flute Sonata in E minor (BWV 1034)	30 bars	O
Cara (2018)	O	22	11 more skilled, 11 less-skilled pianists	Post-task classification (performance speed and accuracy)	O	2.71-7.34		1.36 - 4.71		Complexity	Double-stave music	An excerpt from Ligeti's Etude No. 4, "Fanfares"	27 bars	X

Study	Eye Tracker	Participant			Skill Effect	EHS				Musical Variable	Sight-Reading Material			Playing Tempo Control
		N	Type	Skill Level Classification		Note	Time	Beat	ETS		Type	Source	Length	
Huovinen et al. (2018)	O	37 (exp.1), 14 (exp.2)	14 professional, 23 amateur musicians (exp.1), 14 professional musicians (exp.2)	Pre-task classification (musical background)	O (only for exp.1)				2.12 (exp.1), 2.93 (exp.2) beats	Local complexity (melodic skip)	Single-line melody	12 newly composed melodies (exp.1), 8 newly composed melodies (exp.2)	5 bars (exp.1), 24 bars (exp.2)	O
Chitalkina et al. (2021)	O	22	Musically experienced performers	No level division	X (*both piano sight-reading and sight-singing)				Only relative values available (ms)	Tonality, local incongruence (melodic)	Single-line melody	“Mary Had a Little Lamb” by Lowell Mason	8 bars	O
Imai-Matsumura & Mutou (2021)	O	41	23 experts, 18 nonexperts	Pre-task classification (piano playing experience)	O (correlation analysis between performance ratings and EHS)			0.66 - 1.96		Difficulty	Double-stave music	2 pieces from Sight Playing Workbook of YAMAHA Music Ability Test System	16 bars (easy), 30 bars (difficult)	△ (tempo given prior to playing)

Study	Eye Tracker	Participant			Skill Effect	EHS				Musical Variable	Sight-Reading Material			Playing Tempo Control
		N	Type	Skill Level Classification		Note	Time	Beat	ETS		Type	Source	Length	
Qi & Adachi (2022)	O	32	Advanced pianists	No level division	O (correlation analysis between performance ratings and EHS	2.86-3.12		9.74-13.75		Modality (major, minor), complexity (intervallic)	Double-stave music	6 unknown pieces by Handel and Johann Ernst Bach	20 bars	X
Cara (2023)	O	22	Active musicians (10 professional, 10 undergraduate pianists)	Post-task classification (performance accuracy and speed)	O	4.13		2.47		Structure	Double-stave music	"Slovak Boy's Dance" by Bartók	54 bars	X
Imai-Matsumura & Mutou (2023)	O	39	Professional, college student pianists	No level division	O (correlation analysis between performance ratings, musical experience, and EHS)			1.10 - 1.49		Difficulty	Double-stave music	2 pieces from Sight Playing Workbook of YAMAHA Music Ability Test System	16 bars (easy), 30 bars (difficult)	X

The criteria used to classify participants' skill levels (i.e., pre-task and post-task classification) in Table 2 are similar to those used in the preceding summary of the eye-tracking literature on sight-reading. In the previous review, Qi and Adachi (2022) was the only study that investigated the correlation between eye-related variables and skill level, particularly in the context of performance errors. However, a relatively greater number of studies has analyzed the correlation between the EHS and performance ratings to determine the skill effect—specifically, how the quality of [sight-reading] performance influences the length of the EHS (Imai-Matsumura & Mutou, 2021, 2023; Qi & Adachi, 2022; Rosemann et al., 2016; Sloboda, 1974). This evidence emphasizes the direct relationship between the EHS and sight-reading expertise.

In the table, "Eye Tracker" indicates the utilization of a contemporary eye-tracking device. All the studies employed an eye tracker, except for Weaver (1943), who used a photographic method, and Sloboda (1974, 1977). For the latter, participants were asked to reproduce musical notations to the extent they could remember after the music presentation slide was removed. "EHS" indicates the standard employed for EHS measurement. Historically, the EHS was calculated using note, beat, and time indices. Each value is represented in the table. Until 1997, the EHS was predominantly measured using note and beat indices. For instance, the number of notes and beats that existed between a visual fixation and the performance of a note were counted (Sloboda, 1974, 1977; Truitt et al., 1997; Weaver, 1943). However, the time index was introduced in 1999, which enabled researchers to calculate the latency between a visual fixation and the execution of a note. From this point forward, most studies assessed the EHS using two indices, either note or beat and time. Huovinen et al. (2018) introduced the concept of the eye-time span (ETS), redefining the idea of looking ahead as a metrical distance between a fixation and the corresponding point of metrical time when a performer first fixes his or her eyes on a score. Although the ETS was quantified using the beat index, it differs from the EHS (beat) in its measurement approach. The ETS excludes the hand component of performers and employs a more continuous measurement style. By contrast, the EHS (beat) includes the hand component—representing the quality of performance—and counts discrete events, such as the number of beats.

Another important aspect discussed in the table is “Playing Tempo Control.” This term indicates whether the performance speed of participants was regulated. In the table, “O” represents consistent playing tempi among participants, “X” denotes inconsistencies, and “∆” signifies that a metronome was provided to participants either before or at the start of the sight-reading task but was removed as they continued their performance. The playing tempi varied significantly among participants. Exceptions can be seen in Chitalkina et al. (2021), Huovinen et al. (2018), Penttinen et al. (2015), Rosemann et al. (2016), and Sloboda (1974) where the playing tempo was meticulously controlled throughout the sight-reading task. Imai-Matsumura and Mutou (2021) and Truitt et al. (1997) offered participants a metronome before sight-reading but did not control their tempo during the actual performance. Furneaux and Land (1999) adopted a similar method, using a metronome for the initial two measures and turning it off thereafter, leading to varied tempi among participants (see also Perra et al., 2021; Puurtinen, 2018; Sheridan et al., 2020 for a more comprehensive review of playing tempo within the EHS framework).

In the EHS literature, two central issues parallel findings from eye-tracking studies on sight-reading. The first concerns understanding how the EHS varies based on top-down factors, particularly sight-reading skills, and the second pertains to identifying its susceptibility to bottom-up factors, such as the nature of the music material.

Caution is advised while examining the influence of skill level, a top-down factor, on the EHS. Each study has variations in participants, experimental conditions, and content, leading to myriads of interpretative possibilities (Perra et al., 2021). A comprehensive discussion of the influence of skill level on the EHS is provided in the next section. Nonetheless, a consistent trend observed across the studies is that better sight-readers tend to have a longer EHS on average (Cara, 2018, 2023; Furneaux & Land, 1999; Gilman & Underwood, 2003; Huovinen et al., 2018; Imai-Matsumura & Mutou, 2021, 2023; Penttinen et al., 2015; Sloboda, 1974; Truitt et al., 1997; Wurtz et al., 2009). This suggests that the higher the sight-reading proficiency, the more the eyes lead the hands on the score, allowing more information to be processed within the same timeframe in the buffer (Sloboda, 1984).

Bottom-up influences on the EHS have been linked to specific musical variables. Factors such as texture (Weaver, 1943), phrase boundaries (Sloboda, 1974, 1977), complexity (e.g., Cara, 2018; Huovinen et al., 2018; Qi & Adachi, 2022; Rosemann et al., 2016; Wurtz et al., 2009), and difficulty (Cara, 2023; Imai-Matsumura & Mutou, 2021, 2023) have been explored in depth. Synthesizing these findings shows that a higher degree of musical variables, such as more complex or difficult music, typically results in a shorter EHS. This implies that increased cognitive load limits working memory capacity, thus influencing the behavior of looking farther ahead, as evidenced by the reduced distance between the eyes and hands during sight-reading.

In summary, this section examined the findings of eye-tracking research on sight-reading, covering eye movements, pupil size, the perceptual span, and their complex interactions with both top-down and bottom-up influences. Recognizing the significant role of the motor component and its interplay with visual processing during the sight-reading process, this section underscored the EHS as an integral measure of sight-reading proficiency. The findings of the EHS literature indicated a general trend of longer EHS correlating with sight-reading skills and variable span sizes based on musical parameters. However, despite these insights, studies have shown discernible limitations, which primarily stem from inconsistencies in controlling essential variables, such as the definition of proficient sight-reading and the nature of sight-reading tasks. The following section addresses these concerns in greater detail, particularly highlighting the gaps and subsequent need for further exploration in the literature. This scrutiny sets the stage for the experimental study undertaken in this dissertation.

2.3. Conceptual Framework for the Study: The Three Domains

Before discussing the gaps and limitations in the EHS research, this section introduces a conceptual framework for the experimental study (Figure 1). This framework is crucial for addressing the gaps and limitations and providing an integrated perspective that has often been overlooked in existing research on the EHS. In the framework, this dissertation divides sight-reading into three domains:

musical, cognitive, and behavioral. It investigates how these domains interact, thus understanding sight-reading proficiency holistically. The domain indicators include musical complexity and playing tempo (musical domain), EHS (cognitive domain), and performance accuracy (behavioral domain). The following discussion unfolds across three facets as the study scrutinizes the interrelations among these domains. As the discussion progresses, the logic underlying the trichotomy will be elaborated.

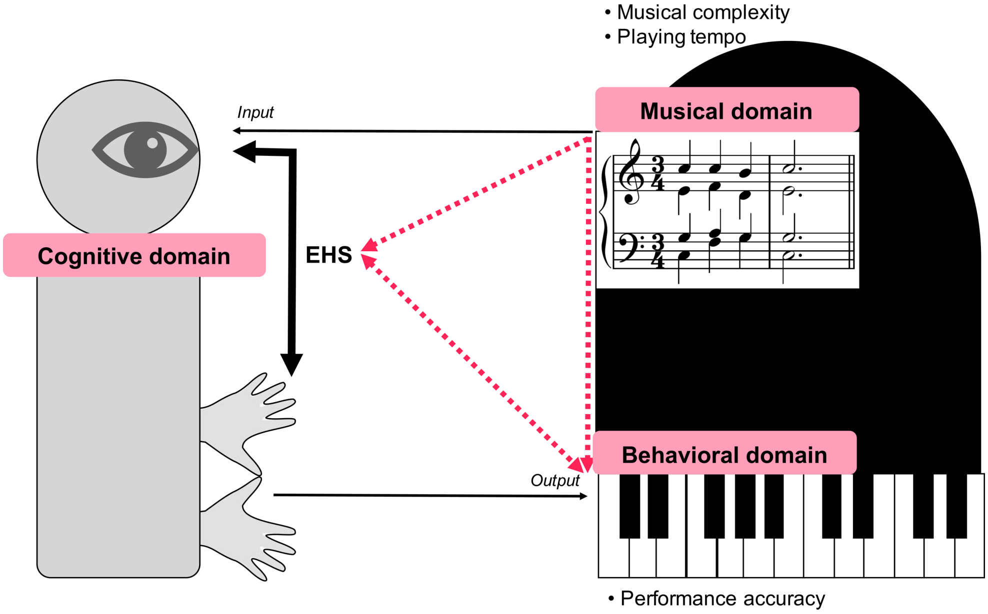

Figure 1. A conceptual framework for the study.

2.3.1. Relationship Between Cognitive and Behavioral Domains

In a pioneering study investigating the relationship between the EHS and sight-reading proficiency, Sloboda (1974) discovered that skilled sight-readers looked approximately 6.8 notes ahead of their hands. By contrast, less-skilled sight-readers looked only 3.8 notes ahead. Thus, the length of the EHS appears to vary based on sight-reading abilities, suggesting that proficient sight-readers can recognize and process musical patterns more efficiently, storing more information in a limited buffer capacity (Furneaux & Land, 1999).

However, a question arises while comparing studies that investigate the direct correlation between the EHS and sight-reading outcomes (Imai-Matsumura & Mutou, 2021, 2023; Qi & Adachi, 2022; Rosemann et al., 2016; Sloboda, 1974). Interestingly, these studies present two contrasting results. Imai-Matsumura and Mutou (2021, 2023) and Sloboda (1974) found a strong positive correlation between the EHS and sight-reading outcomes. By contrast, Qi and Adachi (2022) and Rosemann et al. (2016) observed minimal to no correlation. What could account for this discrepancy? Many studies investigated the relationship between the EHS and sight-reading proficiency, if not focusing solely, then at least also in terms of the quantitative aspects of performance (i.e., the total playing duration). In these studies, participants who played faster were considered skilled sight-readers. Thus, it was inferred that skilled sight-readers, equated with faster players, had a more extended EHS (Cara, 2018, 2023; Furneaux & Land, 1999; Gilman & Underwood, 2003; Imai-Matsumura & Mutou, 2021, 2023; Qi & Adachi, 2022; Truitt et al., 1997; Weaver, 1943; Wurtz et al., 2009).

However, as already pointed out in several studies (Huovinen et al., 2018; Penttinen et al., 2015; Perra et al., 2021; Puurtinen, 2018; Rosemann et al., 2016; Sheridan et al., 2020), a problem associated with uncontrolled playing tempo across participants is that the EHS is supposedly longer while playing the same length of a sight-reading piece at a faster pace. Without a consistent playing tempo among participants, achieving a fair comparison becomes challenging. Variations in this tempo can alter gaze activity in response to sight-reading outcomes, complicating precise analysis. Sight-reading is not merely about getting the notes right. Playing at an acceptable pace is also crucial, as underscored by Lehmann and Kopiez (2016). Sloboda (1977) emphasized this, stating, “It is agreed by musicians that good sight-reading is a remarkable achievement, and the achievement lies in being able to read at anything approaching *the accepted performing speed* [emphasis added] of a difficult piece. Any musician can read perfectly if allowed to go slowly enough” (p. 119). A skilled sight-reader plays with accuracy, and not necessarily speed. In sight-reading, rapid performance without accuracy lacks value. Numerous studies have highlighted the importance of intentionally regulating playing tempo as a key methodological consideration in sight-reading experiments (Huovinen et al., 2018; Lehmann & Ericsson, 1993; Penttinen, 2013; Perra et al., 2022). For instance, Huovinen et al. (2018) noted that knowledge of the influence of controlled playing tempo on looking ahead in sight-reading is limited because research on the EHS has mostly not externally regulated participants’ playing tempo. Therefore, the playing tempo should be controlled across participants, particularly while exploring the relationship between the EHS and sight-reading proficiency.

Only a few studies, such as those by Huovinen et al. (2018) and Penttinen et al. (2015), have examined the relationship between span measurements and sight-reading proficiency while accounting for the qualitative aspects of performance (i.e., performance accuracy) under a strictly controlled playing tempo. The relationship between the EHS and the quality of the performance in not only a single melody but also a double-stave musical piece remains to be explored. Rosemann et al. (2016) is especially noteworthy in this regard for three reasons:

First, their experimental conditions closely mimicked the real sight-reading scenarios with double-stave music that piano sight-readers commonly encountered. Second, the authors objectively measured the EHS by making the participants play at an identical tempo to investigate the correlation between the EHS and sight-reading skills in terms of performance accuracy. Third, all participants were professional pianists. Their findings suggested no significant correlations between the EHS and sight-reading proficiency, implying that a professional's sight-reading strategy may not be limited to looking farther ahead at musical notations (i.e., the eyes are ahead of the hands) as far as possible.

Nevertheless, Rosemann et al. (2016) did not clearly define musical complexity or its characteristics as an independent variable of the EHS. They only considered the EHS in the beat and time indices, excluding the note index. To address these gaps, the present study measures the EHS across note, beat, and time indices at once and investigates the correlations between the EHS and performance accuracy in relation to objectively and quantitatively defined musical complexity.

2.3.2. Relationship Between Musical and Cognitive Domains

A notable observation from the literature is the consistent finding that the average length of the EHS is approximately one second across several studies that used the time index for measurement (Furneaux & Land, 1999; Penttinen et al., 2015; Rosemann et al., 2016; Wurtz et al., 2009). This is intriguing because the EHS values measured in note and beat indices showed variance across studies. Studies that have measured the EHS in the time index have revealed that temporal

EHS remained unaffected by sight-reading skills (Furneaux & Land, 1999) or by the complexity of the music (Wurtz et al., 2009).

How can these findings be reconciled? A plausible interpretation is that sight-readers may consistently read a set distance of musical notations ahead regardless of the number of notes or beats that enter a fixed time window. Perhaps one second is the ideal span for performers to allocate their visual attention on a score and adeptly convert visual information into precise motor actions during a performance (Penttinen et al., 2015; Sheridan et al., 2020). If this holds true, a defining trait of the EHS could be its consistency in time.

However, when the EHS is measured using note and beat indices, the performer's expertise and complexity of the music appear to influence its value. Huovinen et al. (2018) and Penttinen et al. (2015) showed that expert music readers looked farther ahead in scores than did less proficient ones. When these findings are translated to absolute time measures, Huovinen et al. (2018) observed an expertise effect of 300–400 ms in median spans for the more experienced of the two participant groups. As highlighted earlier in the EHS literature review, Huovinen et al. (2018) introduced a novel method to define looking ahead by linking it to the metrical distance—ETS—between fixation and a corresponding point of metrical time at the onset of fixation on the score. Their results indicated that melodic complexity had a significant impact on the ETS.

However, given the absence of studies that simultaneously measure the EHS across all three indices, questions remain: Does the time-consistent EHS hold when all indices are measured concurrently? If it does not, which index provides the most valid representation of the EHS? These questions await further exploration.

2.3.3. Relationship Between Musical and Behavioral Domains

As discussed in the previous section, various factors, including texture, complexity, difficulty, and phrase boundaries, have been extensively researched for their bottom-up influences on the EHS. Among these, complexity is the most frequently explored musical variable (Perra et al., 2021).

Studies have demonstrated that complexity has a significant impact on the EHS. An increase in the global complexity of a musical piece consistently results in a reduction in EHS size (Cara, 2018; Imai-Matsumura & Mutou, 2023; Rosemann et al., 2016; Wurtz et al., 2009). Huovinen et al. (2018) observed that even minor local shifts in relative musical complexity, such as introducing larger melodic intervals, can significantly modify the ETS. Qi and Adachi (2022) further elucidated how complexity interacts with other musical characteristics, such as modality, affecting the EHS. For instance, when the intervallic complexity of a piece in a major modality was reduced, it showed a more extended EHS than did a piece in a minor modality (see also Perra et al., 2021; Sheridan et al., 2020 for a detailed review of the impact of musical complexity on the EHS).

However, there are two primary limitations. First, the definition or standard of musical complexity remains nebulous. Second, the complexity of music is often subjectively examined, described by uncertain musical characteristics with an ambiguous criterion. For example, the difficulty level of sight-reading materials was not represented (Furneaux & Land, 1999) or was determined by the authors' subjective ratings (Rosemann et al., 2016). The sight-reading compositions, either extracted or newly composed, did not accurately reflect actual sight-reading situations, such as the use of short single-line melodies for keyboard sight-reading (as shown in Chitalkina et al., 2021; Penttinen et al., 2015; Sloboda, 1974, 1977; Truitt et al., 1997, with the first study specifically addressing ETS). Whereas some studies have employed existing musical pieces to simulate genuine sight-reading situations (Cara, 2018, 2023; Furneaux & Land, 1999; Gilman & Underwood, 2003; Imai-Matsumura & Mutou, 2021, 2023; Qi & Adachi, 2022; Rosemann et al., 2016; Wurtz et al., 2009), they often lacked clarity in defining musical features such as complexity or did not offer an objective evaluation of these features (Cara, 2018, 2023; Imai-Matsumura & Mutou, 2021, 2023; Rosemann et al., 2016; Wurtz et al., 2009). The EHS was measured under an inequitable condition in which different sight-reading pieces were assigned to different participants (Furneaux & Land, 1999).

Madell and Hébert (2008), in their review of music reading and eye movement, identified two major deficiencies in sight-reading materials: The lack of concentration on the musical structure and detachment from general theories of music perception and cognition. They emphasized the need for a fine-grained approach to assessing musical stimuli, advocating for a focus beyond the "coarsely defined properties" of music, such as well-formed melodies. In this study, two levels of musical complexity are precisely defined in terms of pitch-class distribution and the number of notes per beat. This fine-grained approach to sight-reading materials is expected to offer a lucid exploration of the degree to which musical complexity influences performance accuracy and the length of the EHS.

2.3.4. Aims of the Study

In light of the prevailing gaps and ambiguities in the existing EHS literature on sight-reading, this study aims to comprehensively examine sight-reading proficiency across the musical, cognitive, and behavioral domains. This study endeavors to:

1. Investigate the correlation between the EHS and sight-reading proficiency, particularly in the context of controlled playing tempo and performance accuracy. This relationship is analyzed within the context of double-stave musical pieces.
2. Explore consistent findings across different studies, where the average length of the EHS is estimated to be one second while using the time index for measurement, and discern whether the EHS maintains this consistency across all three indices (note, beat, and time) when measured concurrently.
3. Examine the effects of musical complexity on the EHS and sight-reading proficiency while seeking a more standardized and objective definition of musical complexity. By doing so, this study contributes to a better understanding of the specific aspects and degrees of complexity that influence the EHS and sight-reading proficiency.

By addressing these objectives, this study aims to elucidate the dynamics among the three domains of sight-reading and provide insights that are beneficial for refining methodologies, enhancing interpretative frameworks, and shaping the future direction of sight-reading research.

The experimental study was conducted in collaboration with members of the Music and the Body, an interdisciplinary research project at Seoul National

University. Part of the study was published in 2019 in *Scientific Reports* (Lim et al., 2019), supported by a Seoul National University Research Grant in 2018.

Chapter 3. Methods

3.1. Participants

In all, 31 undergraduate students (30 females and 1 male; mean age = 21.9 years, SD = 2.0 years) majoring in classical piano at Seoul National University participated in this study. They began their piano lessons at an average age of 6 years (SD = 1.0 years, range 4–7 years) and had been playing for an average of 15.9 years (SD = 2.2 years, range 12–20 years). They began majoring in piano with it as their primary instrument at an average age of 11.4 years (SD = 2.3 years, range 6–16 years) and had pursued this major for 10.5 years (SD = 2.9 years, range 5–18 years). Given the time that each participant had dedicated to expertly playing the piano, they were all deemed to be professional pianists in terms of general pianistic skills. No participants were at a novice or intermediate level.

All were right-handed and had either normal or corrected-to-normal vision. The study strictly followed the experimental protocols approved by the Institutional Review Board (IRB) of Seoul National University (1805/003-017). All methods were carried out in accordance with the relevant guidelines and regulations. Written informed consent was secured from every participant. The research was conducted in line with the ethical standards of the 1964 Declaration of Helsinki.

3.2. Sight-Reading Materials

3.2.1. Standards of Musical Complexity

As pitch and rhythm are the two major dimensions of music (Krumhansl, 2000; McAdams, 1989; Schön & Besson, 2002), these elements have been primarily used to modify the complexity or difficulty of the music in the literature (e.g., tonality in Cara, 2018; Waters & Underwood, 1998; note duration in Penttinen et al., 2015; Wurtz et al., 2009; the number of notes in Cara, 2018). Therefore, this study used the degree of pitch-class distribution and the number of notes per beat as determinants of musical complexity. Specifically, the pitch complexity was modulated by the number of accidentals, and the rhythmic complexity was modulated by the number of notes per beat, simultaneous occurrences of two voices per beat, and syncopated notes. For example, a greater presence of chromatic notes (i.e., non-diatonic notes) indicated a higher level of musical complexity. A higher number of notes per beat, simultaneous occurrences of two voices per beat, and syncopated notes also signified increased musical complexity.

3.2.2. Composition

Four musical pieces, each categorized under one of two complexity levels (simple and complex), were specifically composed for this study. The complexity levels were objectively differentiated by comparing the sight-reading materials quantitatively. As detailed in Table 3, the simple and complex pieces varied in terms of the number of accidentals, notes per beat, and total number of notes. None of the pieces contained rests. To eliminate the effect of repetition, each complexity level featured two distinct pieces, where they had quantitatively similar musical components. For instance, any two pieces of the same complexity level maintained the same average number of notes per measure. All pieces were consistent in their key (C Major) and time signatures (4/4 meter), and length (16 measures) to mitigate the influence of confounding variables apart from complexity and playing tempo. Figure 2 displays the four pieces used as the sight-reading materials for the study.

Table 3. Quantitative schema of sight-reading materials.

	Sight-Reading Materials		
Complexity Level	Simple	Complex	*t (P* value*)*
Notes per Beat	2.03 ± 0.71	2.97 ± 0.81	-9.83 (0.000) ***
Accidentals (# or ♭)	0	84	
Half Notes	19	Null	
Simultaneous Occurrences of Two Voices per Beat	0.70 ± 0.48	0.96 ± 0.59	-3.83 (0.000) ***
Syncopated Notes	0	28 - 30	
Key	C Major		
Meter	4/4		
Length	16 measures		
Tempo (BPM)	80 for the slow tempo; 104 for the fast tempo		

***$P < .001$.

Note. Significantly more notes per beat were found in the complex (mean ± SD = 2.97 ± 0.81) than in the simple pieces (mean ± SD = 2.97 ± 0.81) after conducting an independent *t* test [$t(254) = -9.83$ ($P < .001$)]. Significantly more simultaneous occurrences of two voices per beat were found in the complex (mean ± SD = 0.96 ± 0.59) than in the simple pieces (mean ± SD = 0.70 ± 0.48) after conducting an independent *t* test [$t(254) = -3.83$ ($P < .001$)]. The complex pieces were more difficult than were the simple ones in terms of note duration because the former had more syncopation (28–30), whereas the latter did not have any syncopation.
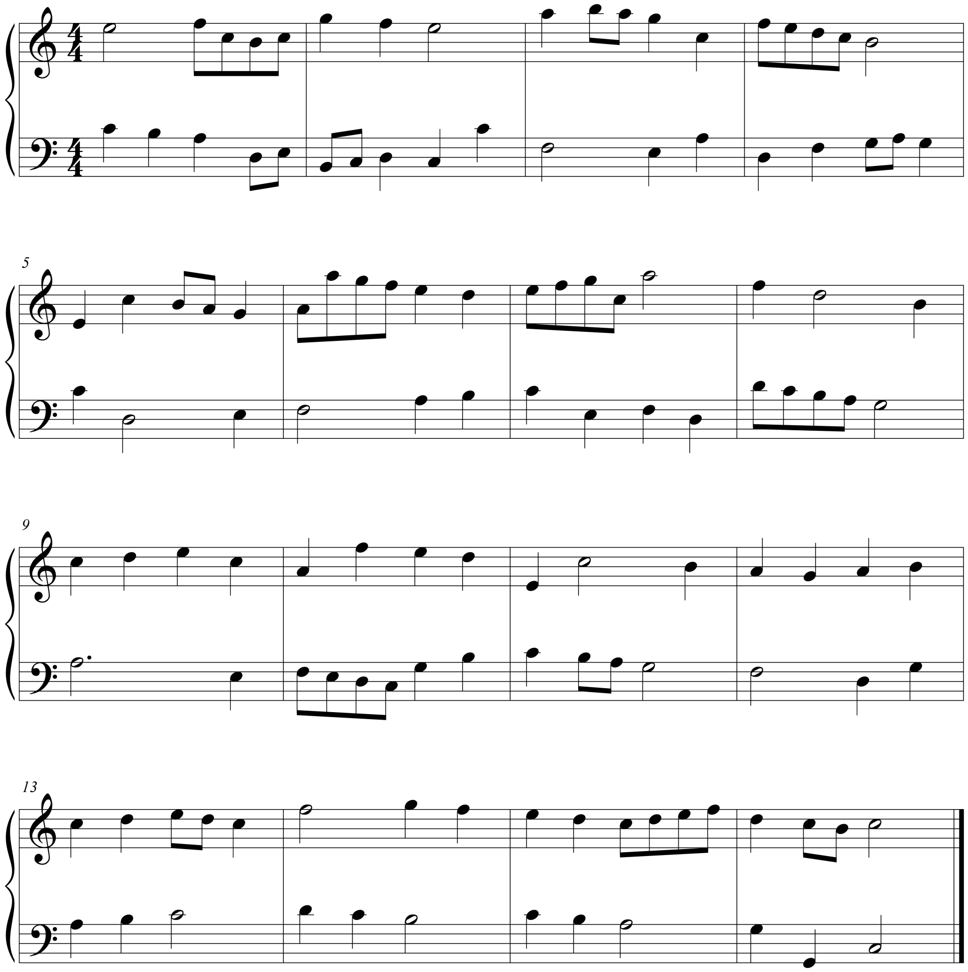
Figure 2.1. Piece 1, simple level.
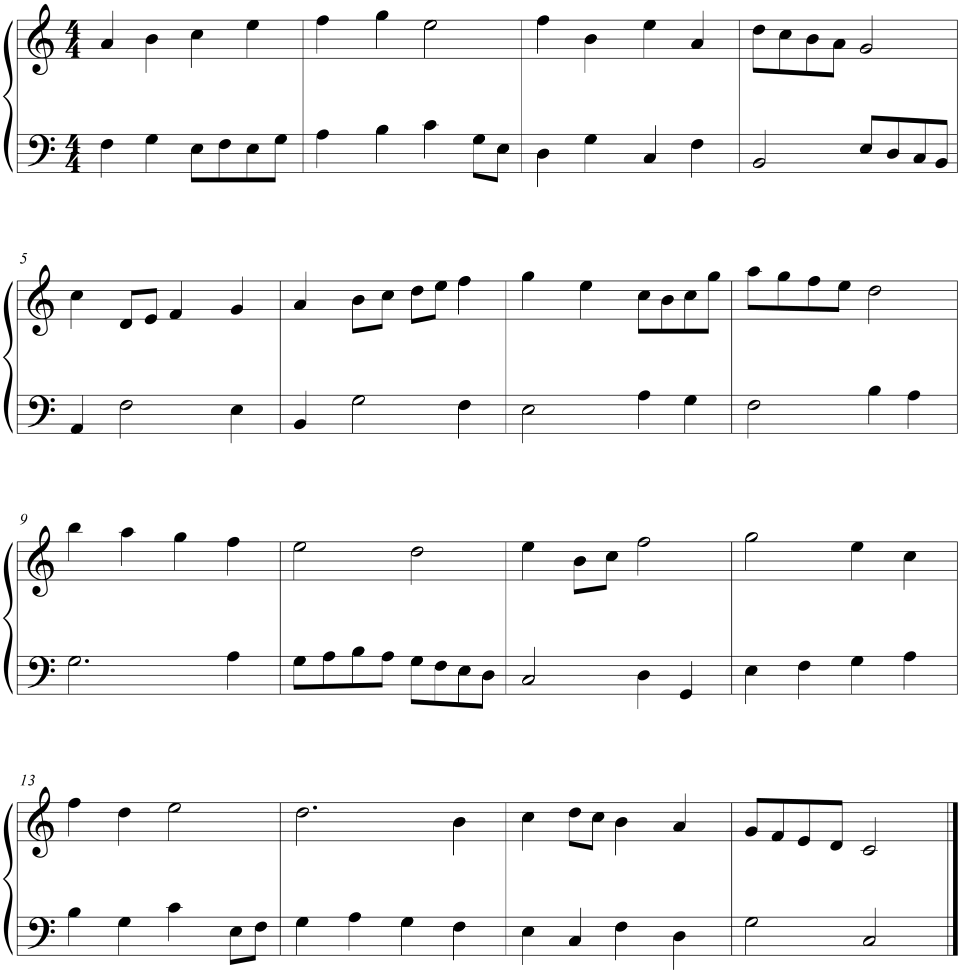
Figure 2.2. Piece 2, simple level.
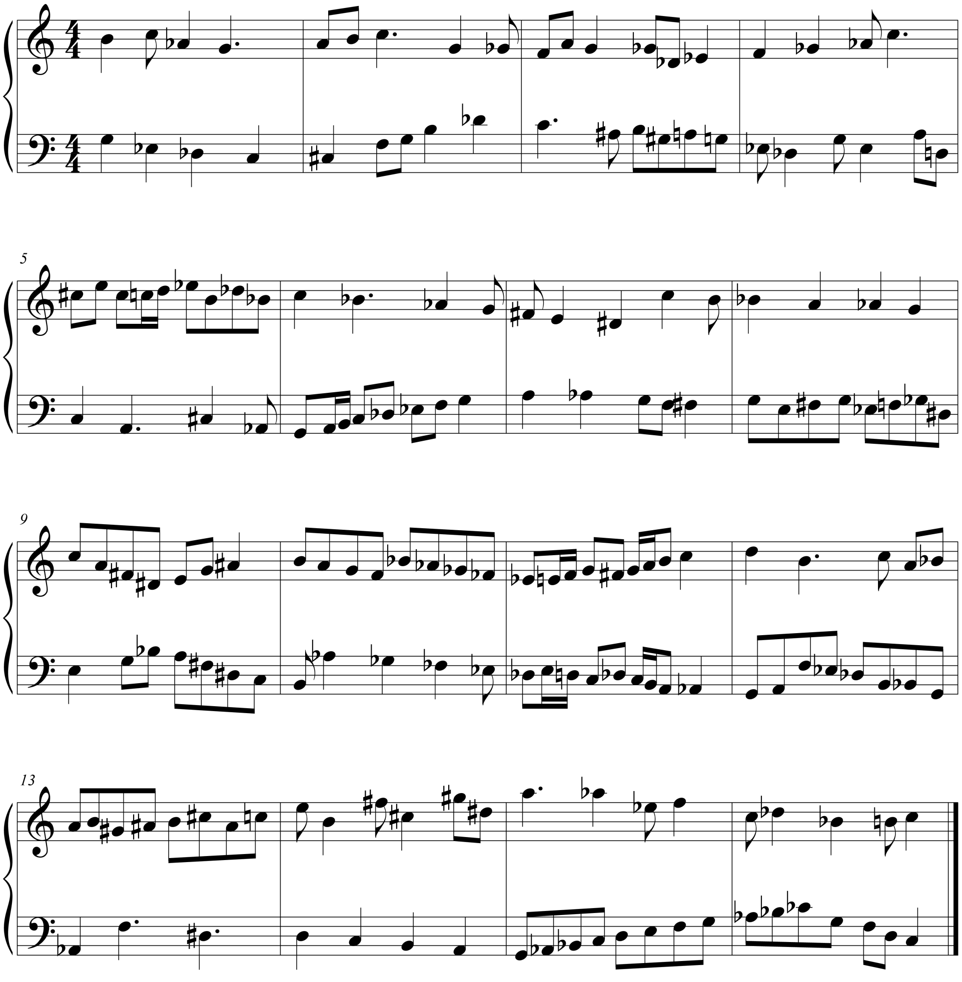
Figure 2.3. Piece 1, complex level.
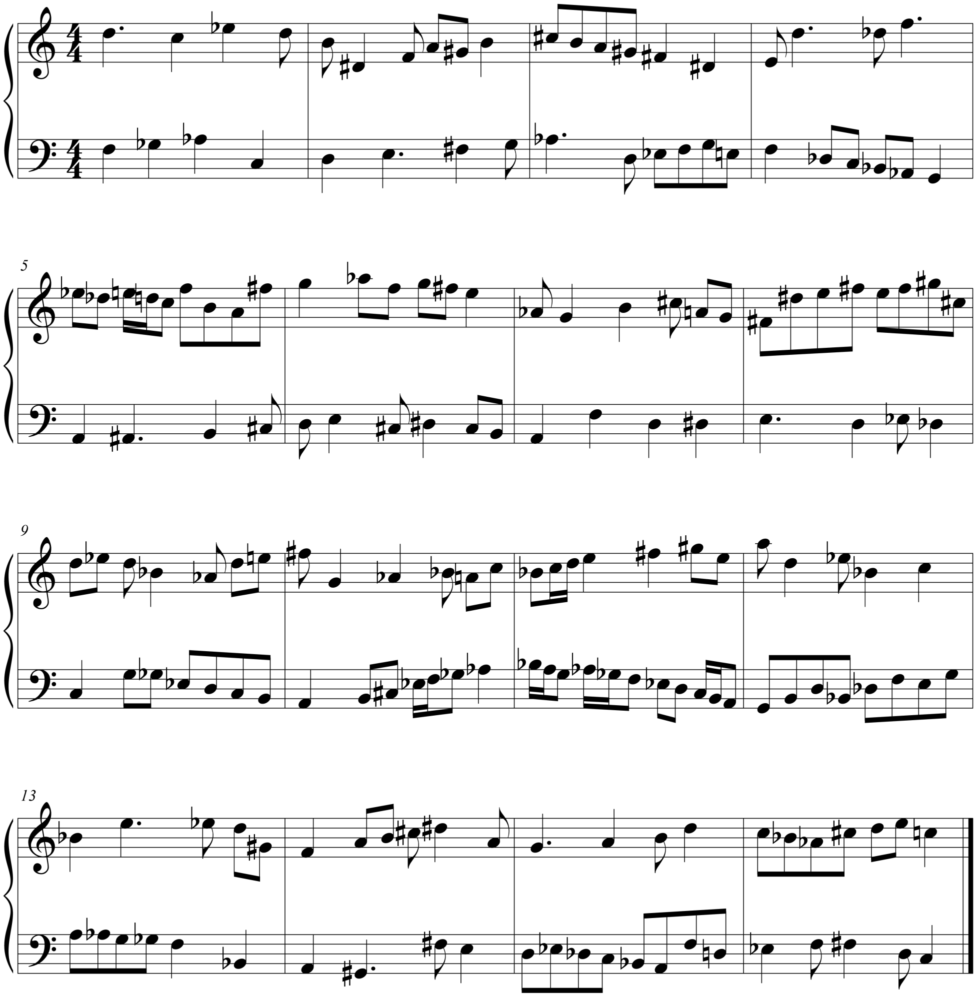
Figure 2.4. Piece 2, complex level.

3.2.3. Measurement of Musical Complexity

The degree of complexity was investigated to demonstrate the nature of the sight-reading materials. The entropy of these materials was calculated and compared to the entropy values of representative composers from various periods (Knopoff & Hutchinson, 1983; Youngblood, 1958). Entropy, stemming from information theory, serves as a mathematical tool to measure the value of information (Shannon & Weaver, 1949). Entropy H is determined by the following formula:

$$\mathrm{H(X)} = -\sum_{i=1}^{n}(P_i \log_2 P_i), \tag{1}$$

where $\mathrm{H(X)}$ represents the entropy of information X, and P_i is the probability of an event occurring with character number i appearing in a stream of characters in Equation (1). Entropy increases as the number of possible outcomes increases and the probability of each outcome becomes equivalent.

In music research, the number of possible outcomes is often equated with the number of pitch classes (Conklin & Witten, 1995; Hiller & Fuller, 1967; Knopoff & Hutchinson, 1983; Temperley, 2007; Youngblood, 1958), with a maximum of 12 probabilities. Entropy becomes higher as all 12 notes appear at an equal frequency. Music employing the 12-tone technique, where all 12 notes are treated with equal importance, is considered the most complex. Therefore, the entropy of music indicates the degree to which a piece is chromatic. To illustrate the level of chromaticism in the sight-reading materials, their entropy was calculated and compared to references from Knopoff and Hutchinson (1983) and Youngblood

(1958). Figure 3 presents a comparison of entropies between the sight-reading materials and references spanning various musical styles. As seen in Figure 3, entropy seems to gradually increase through the annals of Western classical music, starting from the Gregorian chants of the Middle Ages to 20th-century compositions. The simple pieces from this study are positioned between the Middle Ages and the Classical era, whereas the complex pieces exhibit an entropy value closely resembling 12-tone music. Such a comparison facilitates estimating the distribution of pitch classes—how chromatic the sight-reading materials are—compared to the compositions of Western classical music composers.

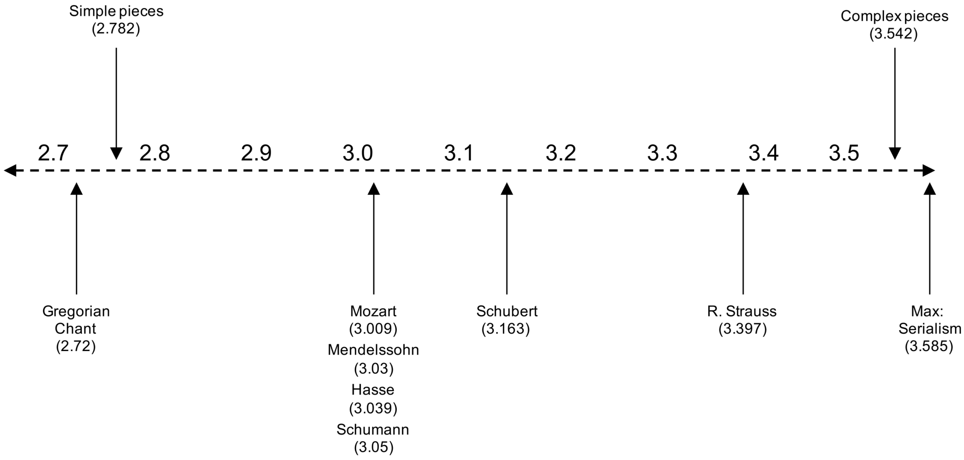

Figure 3. Comparison of the degree of pitch-class distribution between sight-reading materials and compositions of Western classical music composers. The bottom end of the middle-dashed arrow indicates the entropy value of each composer based on the references. For references, this study utilized the entropy values of a Gregorian chant, Mendelssohn, and Schumann from Youngblood (1958), and Mozart, Hasse, Schubert, and R. Strauss from Knopoff and Hutchinson (1983). The upper end represents the entropy value of the sight-reading materials.

3.3. Equipment

The sight-reading materials were displayed on a 23″ monitor with a resolution of 1920 × 1080 pixels. Binocular movements were recorded using Tobii Pro Glasses 2 (Tobii Technology, Stockholm, Sweden) with a sampling rate of 50 Hz (every 20 milliseconds). Participants were positioned 50 cm from the monitor and instructed to keep their heads as stable as possible, but they were permitted to glance at their fingers, thus performing in a natural sight-reading situation. A Yamaha CLP-525 Clavinova digital piano was used in the experiment, and the participants' performances were directly recorded in MIDI format using Logic Pro X 10.2.2.

3.4. Procedures

Each level of complexity had two contrasting tempi. Participants played four sight-reading pieces at two different tempi (simple, slow; simple, fast; complex, slow; and complex, fast). The playing tempi for the two pieces of the same complexity were counterbalanced. Therefore, a piece of a particular complexity could be performed at both slow and fast tempi and was assigned randomly to either tempo for each participant. For instance, if a participant played one of the two pieces of a certain complexity at a fast tempo, they played the other piece from that complexity level at a slow tempo, and vice versa. The designated tempi were 80 BPM for the slow condition and 104 BPM for the fast condition.

The pieces were arranged into eight different sequences for presentation. The sequence did not progressively increase or decrease in complexity, and the playing tempo neither became progressively slower nor faster. This study mirrored the method used by Huovinen et al. (2018) for randomizing presentation orders and assigned the participants to one of the eight sequences by selecting an experimental schedule. This sequence was rotated for each subsequent participant.

Participants were directed to play the sight-reading materials for accurate pitch and rhythm, excluding any musical expression or interpretative elements such as timing, dynamics, or articulation. Before each session began, the eye tracker was calibrated using four distinct points on the sheet music. Participants fixated at each calibration point for a minimum of three seconds. After the calibration phase, a metronome was provided for two measures before playing; then, the participants started to sight-read along with the metronome. The metronome was provided for each beat, assisting participants in maintaining the set tempo. The entire experiment took approximately 20 minutes to complete. Following the main experiment, participants completed a brief questionnaire on their musical background.

3.5. Data Analysis

3.5.1. Eye–Hand Span

Eye movements were recorded with an eye-tracking program (Tobii Glasses Controller by Tobii Technology, Stockholm, Sweden) in real time. After data collection, participants' eye movements were automatically mapped to the corresponding sight-reading score using eye-tracking analysis software (Tobii Pro Lab by Tobii Technology, Stockholm, Sweden).

For each piece, this study created specific events and marked the timestamps for the onset and offset of calibration and sight-reading performance in the raw data. This study only utilized the eye movement intervals between these points for the EHS calculations. To filter and analyze the fixations and saccades, this study applied the Tobii Velocity-Threshold Identification filter, with parameters such as a 75-ms maximum gap length for interpolation, a maximum time/angle between fixations for merging adjacent at 75 ms/0.5 degrees, and discarded fixations below a 60-ms duration (see Olsen, 2012 for more details and rationale behind these parameters). This study determined the onset of fixation as the standard to calculate the EHS length. The analyzed eye movement data were exported in TSV format, and subsequent calculations and visualizations were conducted in the MATLAB program.

The EHS was measured across the four sight-reading pieces. Overall, each sight-reading piece comprised 64 beats, resulting in 64 data points per piece. This study calculated the EHS length for each beat, and the representative EHS for every piece was determined as the mean value derived from the sum of all beat values divided by 64. Specifically, this study calculated the number of notes and beats occurring between the onset of fixation and performance as discrete events and a time delay between two note onsets (latency: sec). The average EHS was calculated by summing all values and dividing by 64 for each index. The beat and time spans were proportional to one another because the time span equaled the beat span multiplied by the playing tempo. For instance, in Figure 4, the green circles indicate the fixation orders corresponding to each beat, represented at the rectangle's upper side. Therefore, the tenth fixation aligns with the first note of beat 12, making the EHS value at this point three notes or two beats.

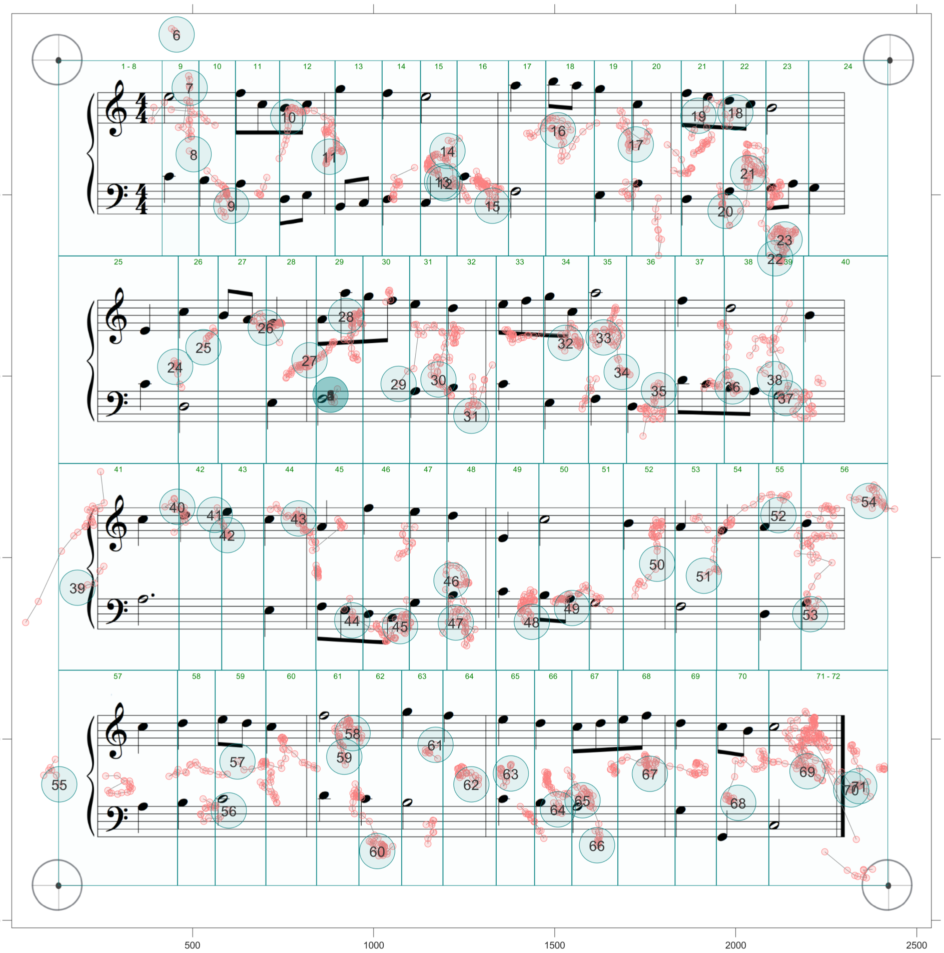

Figure 4. Visualization of the EHS calculation represented as a beat index.

3.5.2. Performance Accuracy

To evaluate performance accuracy, this study examined integrated, pitch, and rhythmic accuracy. Integrated accuracy was analyzed using the dynamic time warping (DTW) algorithm, which compares two distinct sequential datasets differing in time and speed (Müller, 2007; Soulez et al., 2003). In music research, the DTW algorithm has been employed to assess musical performance, providing a quantitative measure of performance accuracy (Bozkurt et al., 2017; Molina et al., 2013; Pan et al., 2017; Vidwans et al., 2017). As it evaluates the similarity between the performance and reference at a frame level (10 ms) rather than at a note level, it can identify subtle differences between both performances. This study used a deadpan MIDI, devoid of variations in dynamics, tempo, or articulation, as the reference performance because participants were instructed to play the sight-reading materials provided accurately only in pitch and rhythm, without adding musical expressions or interpretative elements, to ensure the ecological validity of the investigation. Using the algorithm, this study determined the overall similarities in pitch and temporal details between each performance and the reference MIDIs.

To evaluate performance accuracy both synthetically and individually, this study analyzed pitch and rhythm accuracy. For the former, the number of pitch errors in each performance dataset was counted. Unlike a single melody for which pitch errors can be easily identified manually, the data in this study needed an automated method. The sight-reading materials in this study were polyphonic in style, and with over 124 performance MIDI files (from 31 participants each performing 4 sight-reading pieces), manual counting was impractical. This study used the automated counting method described by Nakamura et al. (2017), which is a state-of-the-art MIDI-to-MIDI alignment method that is available open source. In counting pitch errors, this study considered *missed*, *added*, and *incorrect notes* as types of errors based on Huovinen et al. (2018). *Missed notes* were notes present in the score but omitted in the performance. *Added notes* were extra notes played that were not in the score. *Incorrect notes* were notes played with the wrong pitch. To calculate pitch accuracy, this study divided the total number of pitch errors (sum of the missed, added, and incorrect notes) by the total notes in the piece (130 and 190 for simple and complex pieces, respectively) and then multiplied it by 100 to represent the pitch accuracy percentage.

For rhythmic accuracy, this study counted temporal errors in a manner akin to the approach for pitch accuracy. It computed the inter-onset-interval (IOI) for each reference and performance MIDI and tallied the number of incorrect IOIs in the performance MIDI when compared to the corresponding reference MIDI. An IOI was deemed correct or incorrect based on a predefined threshold. If the deviation of an IOI exceeded this threshold, it was marked as an error, indicating a temporal discrepancy. Taking into account that the shortest note duration in the sight-reading materials was the 16th note and considering variations in timing (ms) playing tempo (whether slow or fast), this study set the threshold at the length of a 32nd note. To calculate rhythmic accuracy, this study divided the number of incorrect IOIs by the total number of IOIs in the piece (129 and 189 for simple and complex pieces, respectively) and then multiplied the result by 100 to express it as a percentage.

Chapter 4. Results

4.1. Performance Accuracy Based on Musical Complexity and Playing Tempo

The integrated, pitch, and rhythmic accuracy based on the four types of sight-reading tasks (simple-slow, simple-fast, complex-slow, and complex-fast) were assessed using a repeated-measures two-way analysis of variance (ANOVA) with musical complexity and playing tempo as factors. For integrated accuracy, this study found that the accuracy values for the simple piece were significantly higher than those for the complex one [F (1, 30) = 314.86, $P < 0.001$] (Tables 4 and 5 and Figure 5a). However, there was no significant difference because of playing tempo [F (1, 30) = 2.38, $P = 0.461$], and there was no interaction effect (complexity × tempo) [F (1, 30) = 1.80, $P = 0.823$]. The pitch and rhythmic accuracy values for the simple piece were significantly higher than those for the complex one [F (1, 30) = 149.16, $P < 0.001$; F (1, 30) = 112.95, $P < 0.001$] and for the slow piece were significantly higher than those for the fast one [F (1, 30) = 68.89, $P < 0.001$; F (1, 30) = 16.31, $P < 0.001$]. There was also an interaction effect (complexity × tempo) [F (1, 30) = 54.04, $P < 0.001$; F (1, 30) = 4.36, $P = 0.045$].

As seen in Figure 5a, the interaction effect suggests that playing tempo had a greater influence on performance accuracy for the complex piece than for the simple piece. To investigate the influence of the four types of sight-reading tasks (simple-slow, simple-fast, complex-slow, and complex-fast) on pitch and rhythmic accuracy, this study performed a repeated-measures one-way ANOVA using the difficulty of the sight-reading tasks as a factor. Pitch and rhythmic accuracy significantly differed based on the difficulty of the sight-reading tasks [F (3, 90) = 125.16, $P < 0.001$; F (3, 90) = 62.06, $P < 0.001$]. Using Bonferroni correction, this study found that the greater the difficulty of the sight-reading tasks, the lower the pitch and rhythmic accuracy (simple-slow > simple-fast > complex-slow > complex-fast; $Ps \leq 0.002$; $Ps \leq 0.02$).

When comparing integrated accuracy measured for the participants' pairs of performances in a given tempo, this study observed a significant positive Pearson correlation between the measurements for the two levels of complexity ($r = 0.48$, $P < 0.001$; Figure 5b). These results show that the participants who played the simple piece in a given tempo relatively accurately also played the complex piece in this tempo with high accuracy. To examine the correlation between the pitch and rhythmic accuracy, a Spearman correlation coefficient analysis was conducted. As demonstrated in Figure 5c, this study found a significant positive correlation between the pitch and rhythmic accuracy (rho = 0.88, $P < 0.001$).

Table 4. Integrated, pitch, and rhythmic accuracy based on musical complexity and playing tempo (mean ± SD).

		Slow	Fast
Integrated Accuracy	Simple	11.92 ± 3.34	12.44 ± 3.60
	Complex	3.47 ± 1.54	3.50 ± 1.85
Pitch Accuracy	Simple	98.66 ± 2.49	96.6 ± 3.46
	Complex	76.67 ± 12.32	60.31 ± 18.58
Rhythmic Accuracy	Simple	99.25 ± 1.3	98 ± 1.08
	Complex	90.72 ± 6.02	86.72 ± 7.61

Table 5. *F* and *P* values of integrated accuracy with different musical complexities and playing tempi.

Factors	*F* value	*P* value
Complexity	314.86	0.000***
Playing Tempo	2.38	0.461
Complexity × Playing Tempo	1.80	0.823

***$P < .001$.

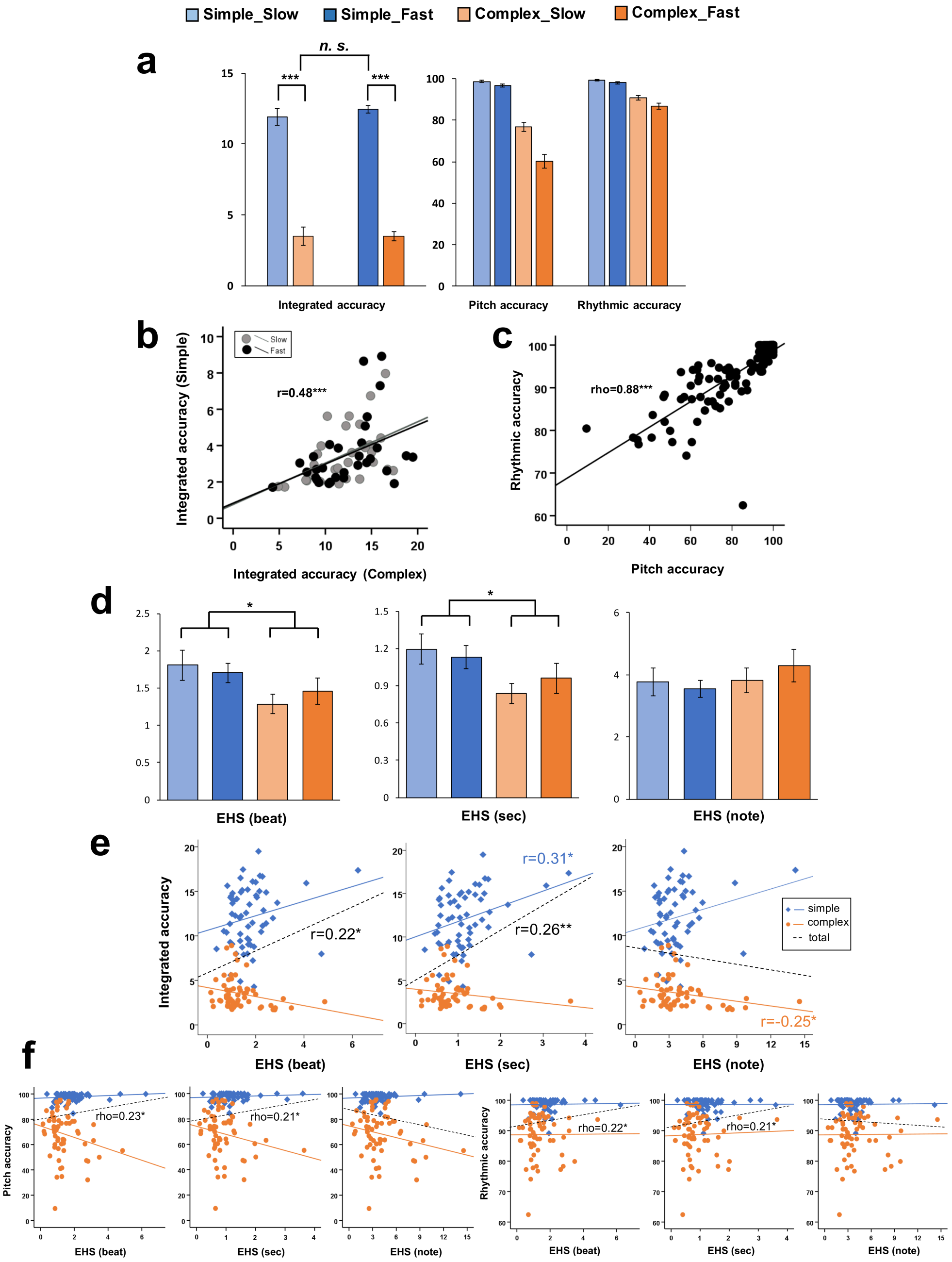

Figure 5. (a) Integrated, pitch, and rhythmic accuracy for different sight-reading tasks (simple-slow, simple-fast, complex-slow, and complex-fast). (b) Scatter plot depicting the regression of correlation between performances of simple and complex pieces in slow and fast tempi in terms of integrated accuracy. (c) Scatter plot depicting the regression of correlation between pitch and rhythmic accuracy. (d) EHS (beat, sec, and note) values for different sight-reading tasks (simple-slow, simple-fast, complex-slow, and complex-fast). (e) Scatter plot depicting the regression of correlation between the EHS (beat, sec, and note) and integrated

accuracy. (f) Scatter plot depicting the regression of correlation between the EHS (beat, sec, and note) and pitch accuracy and correlation between the EHS (beat, sec, and note) and rhythmic accuracy. $^{***}P < 0.001$, $^{**}P < 0.01$, $^{*}P < 0.05$.

4.2. EHS Based on Musical Complexity and Playing Tempo

The EHS values based on the four types of sight-reading tasks (simple-slow, simple-fast, complex-slow, and complex-fast) were evaluated using a repeated-measures two-way ANOVA with musical complexity and playing tempo as factors. Table 6 shows the three types of EHS values (beat, sec, and note) based on complexity and playing tempo. This study found a main effect of complexity but not playing tempo on the EHS (beat and sec) but not the EHS (note). The EHS (beat and sec) values for the simple piece were greater than those for the complex one [$F\ (1, 30) = 6.39$, $P = 0.017$; $F\ (1, 30) = 7.12$, $P = 0.012$]. However, there was no significant difference owing to playing tempo [$F\ (1, 30) = 0.06$, $P = 0.802$; $F\ (1, 30) = 0.11$, $P = 0.741$], and there was no interaction effect (complexity × tempo) [$F\ (1, 30) = 0.64$, $P = 0.431$; $F\ (1, 30) = 0.68$, $P = 0.416$] (Table 7 and Figure 5d). To investigate the influence of the four types of sight-reading tasks (simple-slow, simple-fast, complex-slow, and complex-fast) on the EHS (beat, sec, and note), this study conducted a repeated-measures one-way ANOVA using the difficulty of the sight-reading tasks as a factor. The results indicated no significant differences in the EHS values (beat, sec, and note) based on the task difficulty [$F\ (3, 90) = 2.36$, $P = 0.077$; $F\ (3, 90) = 2.597$, $P = 0.057$; $F\ (3, 90) = 3.029$, $P = 0.605$].

Table 6. EHS with different musical complexities and playing tempi (mean ± SD).

EHS	Complexity	Slow	Fast
EHS (beat)	Simple	1.62 ± 0.78	1.68 ± 0.72
	Complex	1.27 ± 0.76	1.28 ± 0.69
EHS (sec)	Simple	1.10 ± 0.52	1.11 ± 0.52
	Complex	0.83 ± 0.48	0.82 ± 0.38
EHS (note)	Simple	3.33 ± 1.61	3.52 ± 1.55
	Complex	3.77 ± 2.31	3.76 ± 2.02

Table 7. *F* and *P* values of the EHS with different musical complexities and playing tempi.

EHS	Factors	*F* value	*P* value
EHS (beat)	Complexity	6.39	0.017*
	Playing Tempo	0.06	0.802
	Complexity × Playing Tempo	0.64	0.431
EHS (sec)	Complexity	7.12	0.012*
	Playing Tempo	0.11	0.741
	Complexity × Playing Tempo	0.68	0.416
EHS (note)	Complexity	1.06	0.311
	Playing Tempo	0.14	0.715
	Complexity × Playing Tempo	0.59	0.450

*$P < .001$.

4.3. Correlations Between the EHS and Performance Accuracy

This study conducted a Pearson correlation coefficient analysis of the correlation between the EHS and integrated accuracy and Spearman coefficient analyses of the correlation between the EHS and pitch and rhythmic accuracy. Overall, this study found a significant positive correlation between the EHS and integrated accuracy in the indices of beat ($r = 0.22$, $P = 0.016$) and sec ($r = 0.26$, $P = 0.004$; Figure 5e). As seen in Figure 5f, a significant positive correlation was found between the EHS and pitch and rhythmic accuracy in the indices of beat (rho = 0.23, $P = 0.01$; rho = 0.22, $P = 0.014$) and sec (rho = 0.21, $P = 0.022$; rho = 0.21, $P = 0.025$). However, as illustrated in Table 8 and Figures 5e and 5f, this study observed varying correlation tendencies for the integrated and pitch accuracy based on musical complexity. For the simple piece, the EHS values showed a positive correlation with integrated accuracy [$r = 0.31$, $P = 0.015$ (sec)]. For the complex piece, the EHS values showed a negative correlation with integrated accuracy [$r = -0.25$, $P = 0.049$ (note)]. While the differences in correlation tendencies for pitch accuracy were not significant based on musical complexity, the contrasting patterns suggest that participants may have adopted different performance strategies. This variation can be attributed to the changing perceptual difficulties faced by performers during sight-reading tasks. Consequently, this study grouped participants based on their performance accuracy and undertook a correlation analysis between the EHS and performance accuracy, segmented by the difficulty of the sight-reading tasks for each group.

Table 8. Correlation coefficients between the EHS (beat, sec, and note) values and performance accuracy (integrated, pitch, and rhythmic accuracy) values: Pearson correlation coefficients for integrated accuracy (IA); Spearman correlation coefficients for pitch and rhythmic accuracy (*P* value).

		M (*SD*)	IA	Pitch	Rhythmic
Overall	EHS (beat)	1.56 (0.92)	0.22 (0.016)*	0.23 (0.010)*	0.22 (0.014)*
(*N* = 124)	EHS (sec)	1.03 (0.60)	0.26 (0.004)**	0.21 (0.022)*	0.21 (0.025)*
	EHS (note)	3.86 (2.34)	-0.05 (0.570)	-0.06 (0.523)	-0.05 (0.597)
Simple	EHS (beat)	1.76 (0.94)	0.22 (0.085)	0.12 (0.367)	-0.1 (0.432)
(*n* = 62)	EHS (sec)	0.987 (0.50)	0.31 (0.015)*	0.06 (0.664)	-0.15 (0.248)
	EHS (note)	3.68 (1.93)	0.23 (0.078)	0.11 (0.396)	-0.12 (0.372)
Complex	EHS (beat)	1.37 (0.86)	-0.24 (0.057)	-0.19 (0.133)	0.08 (0.522)
(*n* = 62)	EHS (sec)	0.90 (0.58)	-0.16 (0.207)	-0.19 (0.146)	0.09 (0.487)
	EHS (note)	4.06 (2.59)	-0.25 (0.049)*	-0.18 (0.171)	0.08 (0.515)

***$P < 0.001$, **$P < 0.01$, *$P < 0.05$.

4.4. Correlations Between the EHS and Performance Accuracy Based on Sight-Reading Task Difficulty in High- and Low-Accuracy Groups

This study categorized participants into three groups (*n* = 10, 11, and 10) based on their performance accuracy values for the four types of sight-reading tasks (simple-slow, simple-fast, complex-slow, and complex-fast). The aim of the study was to explore the varying correlation patterns between the EHS and performance accuracy contingent on sight-reading task difficulty. To this end, this study conducted Spearman correlation coefficient analyses between the EHS (beat, sec, and note) values and performance accuracy (integrated accuracy, pitch accuracy, and rhythmic accuracy) values for each group. To clarify the group differences, this study conducted statistical analyses on the top 10 participants with the highest performance accuracy and the bottom 10 participants with the lowest performance accuracy for each task.

Table 9 and Figure 6 illustrate the correlation patterns between the EHS and integrated accuracy as modulated by the difficulty of the sight-reading task. Despite the presence of several outliers, these were addressed as ranks in the nonparametric analyses. For the high-accuracy group, there was a positive correlation between the EHS (beat, sec, and note) and integrated accuracy in the easiest task (simple-slow) [Spearman's rho = 0.75, *P* = 0.013], whereas a negative correlation was evident for the most challenging task (complex-fast) [rho = -0.78, *P* = 0.008 (beat), rho = -0.69, *P* = 0.029 (sec), rho = -0.76, *P* = 0.011 (note)]. By contrast, the low-accuracy group displayed an overarching negative correlation. A notably significant negative correlation emerged for the simple-fast task [rho = -0.72, *P* = 0.019 (beat), rho = -0.71, *P* = 0.022 (note)]. For pitch and rhythmic accuracy, no significant correlations with the EHS were identified relative to the sight-reading task difficulty.

Table 9. Correlations between the EHS and integrated accuracy in high- and low-accuracy groups.

Performance Group	EHS	Integrated accuracy [Spearman's rho (*P* value)]			
		Simple-Slow	Simple-Fast	Complex-Slow	Complex-Fast
High (*n* = 10)	EHS (beat)	0.75 (0.013)*	0.10 (0.777)	0.08 (0.829)	-0.78 (0.008)**
	EHS (sec)	0.75 (0.013)*	0.44 (0.20)	0.07 (0.855)	-0.69 (0.029)*
	EHS (note)	0.75 (0.013)*	0.10 (0.777)	0.08 (0.829)	-0.76 (0.011)*
Low (*n* = 10)	EHS (beat)	0.10 (0.777)	-0.72 (0.019)*	-0.16 (0.651)	-0.61 (0.060)
	EHS (sec)	-0.06 (0.881)	-0.38 (0.276)	-0.07 (0.855)	-0.65 (0.043)*
	EHS (note)	0.10 (0.777)	-0.71 (0.022)*	--0.18 (0.627)	-0.58 (0.082)

**$P < 0.01$, *$P < 0.05$.

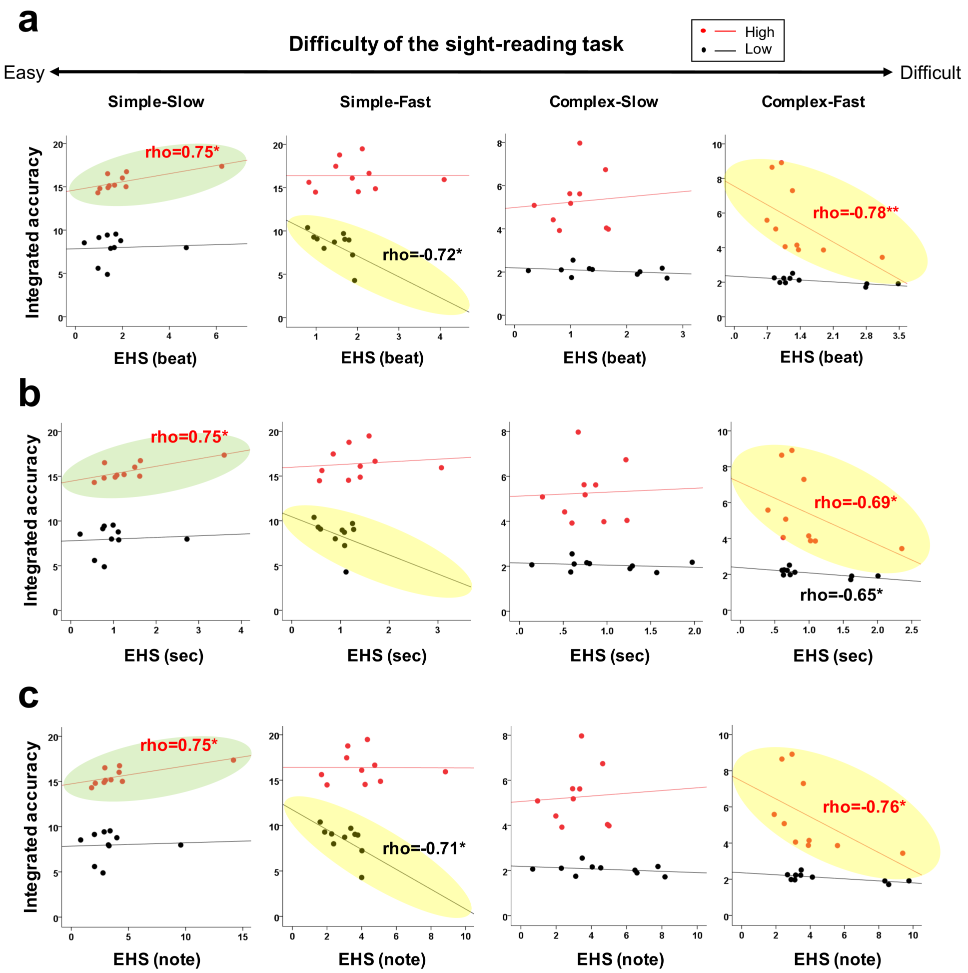

Figure 6. Correlations between the EHS (beat, sec, and note) and integrated accuracy across varying difficulties of the sight-reading task for high- and low-accuracy groups. (a) Correlation between the EHS (beat) and integrated accuracy. (b) Correlation between the EHS (sec) and integrated accuracy. (c) Correlation between the EHS (note) and integrated accuracy. $^{**}P < 0.01$, $^{*}P < 0.05$.

Chapter 5. Discussion

This chapter discusses the primary findings and their implications. The discussion unfolds across three facets of the conceptual framework in parallel with the previous discussion in Chapter 2, namely the relationships between cognitive (EHS) and behavioral (performance accuracy) domains, musical (musical complexity and playing tempo) and cognitive domains, and musical and behavioral domains. This chapter critically evaluates the study's limitations and suggests avenues for future research on the EHS.

5.1. Main Findings and Implications

5.1.1. Relationship Between Cognitive and Behavioral Domains

The most intriguing finding of the present study is the varying relationship between the EHS (cognitive domain) and performance accuracy (behavioral domain) based on the difficulty of the sight-reading tasks (musical domain). Specifically, when categorizing performer groups based on their accuracy for the four sight-reading tasks (simple-slow, simple-fast, complex-slow, and complex-fast), this study observed distinct correlation patterns between the EHS and performance accuracy under different task conditions. For the high-accuracy group, a positive correlation emerged between the EHS (beat, sec, and note) and integrated accuracy in the easiest task (simple-slow), whereas a negative correlation was evident in the most challenging task (complex-fast). On the other hand, the low-accuracy group demonstrated a negative correlation overall in the sight-reading tasks, with a particularly significant negative correlation in the simple-fast task. These findings show that it is difficult to determine sight-reading proficiency, relying solely on the EHS. Whereas the results are compatible with those of Huovinen et al. (2018) and Penttinen et al. (2015), who noted that proficient music readers generally look farther ahead in scores than did their less-skilled counterparts, this study also indicates that the EHS may not be a decisive indicator of sight-reading proficiency. It posits that a proficient sight-reader does not invariably exhibit a longer EHS. Similarly, a less proficient sight-reader does not necessarily manifest a shorter EHS. Instead, this study suggests that the relationship between the EHS and performance accuracy may be strongly influenced by the difficulty of sight-reading tasks.

Participants with above-average sight-reading skills tend to perform more accurately and can look farther ahead at will (resulting in a longer EHS) during relatively easy sight-reading tasks. This study hypothesizes that these varying correlations between the EHS and performance accuracy based on sight-reading proficiency may stem from the different degrees of perceptual difficulty of the performer experiences during sight-reading tasks. Although the combination of musical complexity and playing tempo objectively sets the difficulty level for sight-reading tasks, the perceptual challenges may differ based on an individual’s sight-reading skill. This difference causes varying correlations between the EHS and performance accuracy. For instance, for the high-accuracy group, having a wider EHS in the easiest task (simple-slow)—meaning their eyes looking ahead of their hands while sight-reading—was advantageous to improving performance accuracy. By contrast, in the most challenging task (complex-fast), maintaining a narrower EHS—implying that their eyes do not look too far ahead of the notes being played—benefited their sight-reading outcome. Intriguingly, the low-accuracy group, which showed a negative correlation between the EHS and performance accuracy across all four difficulty levels overall, found that maintaining the narrowest EHS in the moderately challenging task—the simple-fast condition—was most advantageous for improving performance accuracy. This strategy was similar to the one employed by the high-accuracy group in the most difficult task (complex-fast), suggesting that the low-accuracy group may have found the simple-fast task as challenging as the high-accuracy group found the most difficult task. Given that the simple complexity level of sight-reading materials used in this study was not particularly challenging in terms of a general instrumental technical level and considering that all participants were professional pianists who had similar levels of instrumental technique, it seems plausible that the shorter EHS strategy shown by the low-accuracy group in the simple-fast task stemmed from their perceptual difficulties. Therefore, a comprehensive understanding of sight-reading proficiency necessitates evaluating the EHS within a multidimensional framework that includes various musical domains rather than a singular referent point.

Relatively few studies have investigated the interplay of cognitive and behavioral domains in sight-reading, which embodies visuomotor coordination, especially in the musical context. For instance, Huovinen et al. (2018) examined how local modifications in a musical stimulus affected the ETS. Their findings showed that the ETS was significantly influenced by local melodic complexity while sight-reading single-line melodies. Aside from the need to investigate similar complexity effects in piano compositions represented on dual staves, the starting point of this study was a notable discrepancy in the extant research on the relationship between the EHS and sight-reading outcomes. Imai-Matsumura and Mutou (2021, 2023) and Sloboda (1974) observed a strong positive correlation between the EHS and sight-reading outcomes. However, Qi and Adachi (2022) and Rosemann et al. (2016) did not detect such a correlation. This study provides a detailed interpretation of the relationship between cognitive and behavioral domains, factoring in the challenge level of the musical domain. If the EHS is a strategy adaptable to the demands of sight-reading tasks, the study's findings can be in harmony with those of both sets of prior studies.

One potential reason for the contrasting correlations in earlier research may be the varying levels of difficulty of their sight-reading tasks. For example, the sight-reading materials used by Sloboda (1974) corresponded to the easy level of difficulty in this study, leading to a strong positive correlation between the EHS and sight-reading outcome. By contrast, as the sight-reading materials used by Rosemann et al. (2016) ranged between the easy and difficult levels of this study, a significant positive correlation was not found between the EHS and sight-reading outcome. Research has suggested that skilled sight-readers may have a heightened sensitivity to the features of music. For instance, Sloboda (1974, 1977) found that the EHS often extends to phrase boundaries, especially among accurate sight-readers. Penttinen et al. (2015) noted that experienced performers often had a shorter EHS while playing altered melodies. While not based on skill level, the EHS size within individual performers varied based on the given sight-reading piece. Weaver (1943) emphasized this variability, stating that participants exhibited differing EHS sizes according to the types of music, which included various musical textures such as harmonic, melodic, or accompanied-melody music. Some participants, for instance, had a shorter span for harmonic music but a longer one for melodic music. This adaptability in the EHS was consistent, with no evidence of the span remaining static for over two consecutive measures. The present study draws from these studies and infers that the musical domain can moderate expertise-driven differences in cognitive processing. The findings offer empirical evidence on this perspective, supporting the importance of a flexible EHS as a sight-reading strategy, and suggest that proficient sight-readers do not necessarily maintain a longer EHS than do their less-skilled counterparts. Instead, they adjust their cognitive strategies flexibly in response to moderating factors such as the characteristics of the music.

5.1.2. Relationship Between Musical and Cognitive Domains

This study investigated the effect of musical complexity and playing tempo (musical domain) on the EHS (cognitive domain). It determined that musical complexity significantly influenced the EHS. Interestingly, the value of the EHS (beat and sec) during the performance of a complex piece was smaller than that during a simple piece. However, the value of the EHS (note) remained consistent irrespective of the musical complexity. The data suggested that the EHS (beat and sec) did not significantly change according to the playing tempo (whether slower or faster), whereas several studies have shown the significant effect of the playing tempo on the EHS in the time index (Furneaux & Land, 1999; Wurtz et al., 2009). A striking observation from prior studies is that the influence of musical complexity on the EHS has been shown only for a beat (Cara, 2018; Huovinen et al., 2018; Rosemann et al., 2016) or note index (Cara, 2018; Wurtz et al., 2009), and the EHS in the time index remained unaffected by musical complexity (Furneaux & Land, 1999; Rosemann et al., 2016; Wurtz et al., 2009). By contrast, in this study, the EHS (beat and sec) were affected by musical complexity, with no observable difference in the EHS (note) based on musical complexity. The findings imply that musical complexity may alter the efficiency of processing musical notations in a limited buffer capacity rather than confirming the existence of a consistent time lag between eye and hand movements.

This conclusion is consistent with Huovinen et al. (2018), who found that the complexity of the upcoming symbols of a score affects saccadic processes, although the tendency of the influence of musical complexity differed between Huovinen et al. (2018) and the current study. In this study, performers had a shorter span when presented with the complex sight-reading material, whereas Huovinen et al. (2018) observed that the span lengthened in response to musical complexity. The contrasting tendency in the effect of musical complexity on the span may be because of the different types of approaches used for the span measurements (Huovinen et al., 2018). In Huovinen et al. (2018), musical complexity referred to a local phenomenon, such as a large intervallic leap, and the authors aimed to determine if the span adjusts locally in these cases, leading them to introduce the Early Attraction Hypothesis. This hypothesis posits that when musicians encounter visually or musically complex notes while sight-reading a score, their eyes are drawn to these notes sooner than they are to simpler notes. The hypothesis implies that in anticipation of forthcoming complex note symbols or relationships in the musical score, the reader’s span of looking ahead expands. These complexities capture the reader’s attention sooner, helping allocate sufficient processing time to navigate these challenges. However, in the present study, the complexity was not just localized, but also encompassed entire music pieces that can be viewed as global complexity. Some pieces were intrinsically more complex than others, and the span measurements taken were in alignment with these overall distinctions in complexity. Thus, the findings from Huovinen et al. (2018) and the current study could be true simultaneously: (1) Compared to simple pieces, in more complex ones, musicians use somewhat shorter spans in general, and (2) when there are local differences in musical complexity, more complex elements may attract the reader's eyes earlier, yielding locally longer spans.

Why did the EHS (beat and sec) vary with musical complexity, while the EHS (note) remained unchanged? This study surmises that this variation arises from the varying number of notes per beat between the two complexities. The complex piece contained significantly more notes per beat than did the simple piece, leading to two potential explanations. One possibility is that the EHS may be note-consistent independent of musical complexity. As it followed a constant number of notes, the length of the EHS (note) may have been different in the two complexities. However, the EHS (beat and time) may have varied because of the visual density of the presented notes, although both pieces comprised the same number of notes. Another possibility is that the EHS may differ based on musical complexity, suggesting that the length of the EHS changed because of the characteristics of musical variables. For instance, if musical complexity is defined in terms of other musical elements instead of the number of notes, the value of the EHS will differ. Nonetheless, this outcome is irrelevant to sight-reading proficiency because performance accuracy was not included in the result.

5.1.3. Relationship Between Musical and Behavioral Domains

This study investigated the influence of musical complexity and playing tempo (musical domain) on performance accuracy (behavioral domain) and discovered that while musical complexity significantly affected performance accuracy, playing tempo did not. It offered more concrete evidence of the influence of musical complexity. Although studies on the EHS have indicated that musical complexity significantly affects sight-reading outcomes (Cara, 2018; Imai-Matsumura & Mutou, 2023; Rosemann et al., 2016; Wurtz et al., 2009), a limited understanding remains of the specific types of complexity and the degree to which they affect sight-reading. To elucidate the influence of musical complexity objectively and concretely, this study demonstrated the degree of pitch-class distribution by comparing the entropy of sight-reading pieces with references from various Western classical music composers, thus representing how chromatic the sight-reading pieces are. This study demonstrated that the complexity level of simple pieces is comparable to that of music from the Middle Ages to the Classical period, whereas the complexity level of complex pieces resembles that of 12-tone music. Using a quantitative approach to assess the qualitative aspects of musical complexity enabled the establishment of a more apparent threshold for the influence of musical complexity on the EHS and sight-reading proficiency.

As expected, this study found a strong positive correlation between participants' performance accuracy for both simple and complex pieces. Those who excelled with simple pieces also did well with complex ones. This indicates that the performances of proficient sight-readers remain consistent regardless of musical complexity, at least based on the complexity parameters of this study. If this is true, the following question arises: Who are these skilled sight-readers, and how do they maintain such consistency? These questions extend beyond the purview of this study. However, previous studies have offered some insights. As mentioned in Chapter 2, studies have emphasized both practice-dependent and practice-independent factors that influence sight-reading proficiency (Ericsson et al., 1993; Kopiez & Lee, 2008; Meinz & Hambrick, 2010; Platz et al., 2014). Kopiez and Lee (2006) noted that the combination of predictors for successful sight-reading changes based on the complexity of the sight-reading material. For simpler pieces, general pianistic expertise, which is linked to experience and practice, is the primary determinant of superior sight-reading. For the most complex pieces, psychomotor speed, associated with inherent cognitive abilities, becomes the primary predictor. Considering this study's results, both practice-dependent and practice-independent factors seem to have a proportional relationship or at the very least a positive correlation, as performers who excelled in the simple pieces did the same for the complex ones. The correlation between the EHS and performance accuracy in the high-accuracy group showed that for the easiest and most difficult tasks, a longer and shorter EHS were linked to accurate performance, respectively. This may imply that practice-dependent factors are related to a longer EHS, with eyes reading scores farther ahead of hand movements, whereas practice-independent factors may correspond to a shorter EHS, where eyes and hands move nearly in sync. However, this is only a tentative hypothesis without any empirical evidence. Detailed research exploring this theory is warranted.

In conclusion, this study explored the interrelationships among the three domains of sight-reading, namely musical complexity and playing tempo (musical domain), EHS (cognitive domain), and performance accuracy (behavioral domain). The findings of the varying correlations between the EHS and performance accuracy depending on the difficulty of sight-reading tasks suggest that 1) the EHS is not a decisive indicator of sight-reading proficiency but is a strategy that can be changed according to moderating factors of the musical domain, such as musical complexity, and 2) proficient sight-readers are performers who are skilled in adjusting their EHS instead of always maintaining an extended span.

5.2. Limitations and Future Directions

In this study, the statistical significance of the correlation between the EHS and performance accuracy varied based on the specific accuracy measurement used. While dividing the participants into three groups based on their accuracy values and investigating the correlation between the EHS and performance accuracy for the four sight-reading tasks, this study found significant correlations, whether positive or negative, of the EHS with integrated accuracy based on the task difficulty. However, there was no significant correlation between the EHS and either pitch or rhythmic accuracy. Why did the relationship between the EHS (cognitive domain) and performance accuracy (behavioral domain) change according to the accuracy measurements?

One possible explanation is the intrinsic differences in evaluation methods. Integrated accuracy, which assesses all performance elements collectively, may offer a more holistic perspective on the quality of sight-reading outcomes compared to individual accuracy, such as pitch and rhythmic accuracy. Whereas individual accuracy metrics for pitch and rhythm were not strongly tied to the EHS, evaluating these elements together mat have revealed a unique relationship with the EHS. Another factor to consider is the nature of the algorithm used for performance evaluation. This study employed the DTW algorithm to evaluate integrated accuracy, setting a deadpan MIDI as a reference and comparing it with participants' performances. The principle is how closely the participants played in comparison to the reference performance. This approach was feasible because this study instructed participants to play only accurately, excluding any expressive elements. However, typical MIDI data include information not only about pitch (i.e., whether the correct or incorrect note was hit) and rhythm (i.e., onset and offset timing of a note) but also about velocity (i.e., the force with which a note hit). Thus, even though this study instructed participants to perform accurately without any expressiveness, considering that all participants were professional pianists, they may have inadvertently introduced expressiveness into their performances, which could be reflected in velocity while calculating the overall similarity matrix. This may explain why the correlation between the EHS and performance accuracy differed between integrated accuracy and individual (pitch and rhythmic) accuracy.

This study suggests several avenues for future research on the EHS. First, when investigating the correlation between the EHS and sight-reading proficiency, expressiveness should be taken into account along with accuracy for performance evaluation. The current study evaluated sight-reading outcomes solely in terms of accuracy, excluding musical expressions such as dynamics, articulation, or expressive timing. Such an evaluation was primarily to maintain the ecological validity of the sight-reading experiment. Rigorously controlling their performance conditions was crucial to prevent the potential influence of unexpected performance variables and ensure a fair comparison of sight-reading outcomes among participants. For instance, this study made all participants perform at the same tempo and required them to perform only accurately without including any musical expression. This is because if there are deviations other than accuracy, the comparative value among the performances could diminish, and the resulting findings may be skewed. However, although accurate performance is a foundational requirement for proficient sight-reading, a performance adorned with abundant expression and interpretation, similar to a well-rehearsed performance, can ultimately be considered the highest level of sight-reading. As emphasized in the definition of sight-reading outlined in Chapter 1, the ultimate goal of sight-reading is the execution of musical passages with the appropriate and natural musical flow and expressivity. Therefore, if future research considers not only accuracy but also expressive aspects while evaluating performance outcomes and investigates the relationship between the EHS and sight-reading proficiency, doing so can genuinely contribute to unraveling superior sight-reading abilities.

To date, few studies have evaluated participants' sight-reading outcomes, including musical expression, while investigating the relationship between the EHS and sight-reading proficiency (Imai-Matsumura & Mutou, 2021, 2023; Rosemann et al., 2016). However, these studies did not clearly describe which expressive elements were employed as evaluation indicators. Accuracy and expressiveness were not evaluated separately, but were incorporated into a single assessment (e.g., rating components were the number of mistakes in sound, beat, and rhythm, the quality and accent of sound, and playing tempo in Imai-Matsumura & Mutou, 2021, 2023; the number of mistakes, omitted notes, and correct phrasing in Rosemann et al., 2016). However, combining performance accuracy (a quantitative aspect) with expressiveness (a qualitative aspect) into one metric may make it challenging to determine whether a particular research outcome is because of either accurate or expressive performance. If future research assesses accuracy and expressiveness individually and explores the relationship of each performance outcome with the EHS and/or musical variables, it would be possible to better understand sight-reading proficiency from musical and scientific perspectives. One crucial point is that expressiveness, such as performance accuracy, should be quantifiable. Only then can we better explain which aspects of musical expression (e.g., deviations in timing or dynamics) are related to the flexible adjustment of the EHS. There has already been considerable research on quantifying the expressiveness of performance (Cook, 2009; Repp, 1990, 1998, 1999; Todd, 1992; see also Part 3 of Fabian et al., 2014 for a systematic review). If research on the EHS were to more fully embrace such quantitative approaches to evaluating expressive sight-reading, it would significantly contribute to a more sophisticated explanation of the complexities of genuinely skilled sight-reading.

Second, this study recommends further exploration of the correlation between the EHS and sight-reading proficiency across a broader range of musical styles. It indicates that the EHS is not fixed but dynamic, changing based on the properties of the sight-reading piece, specifically task difficulty, which is determined by musical complexity and playing tempo. More proficient sight-readers seem to better utilize this flexibility. Participants who excelled in sight-reading tended to use a longer EHS strategy, allowing their eyes to look farther ahead of their playing hand during more manageable tasks. By contrast, for more challenging tasks, they favored a shorter EHS, where their reading and playing actions were almost synchronized.

However, these strategic patterns may differ based on various musical styles, other than task difficulty determined by musical complexity and playing tempo. Considering the actual sight-reading situations of musicians, it becomes evident that they perform a wide variety of pieces from different eras and styles at first sight. For instance, in the process of building their repertoire, pianists sight-read pieces ranging from piano sonatas by Franz Joseph Haydn (1731–1809) to character pieces by Robert Schumann (1810–1856), and even piano etudes by György Ligeti (1923–2006). This is because every excellent rehearsed performance begins with reading the score. Furthermore, the cognitive strategies performers employ while sight-reading these different styles of music, whether conscious or not, are likely to vary. For example, pattern recognition would be more utilized while sight-reading Classical music, where conventional rhythmic patterns such as arpeggios are widely used. This can result in a more extended EHS, as pianists group notes into familiar patterns, thus reducing cognitive load and allowing for a time lag that lets the eyes lead significantly ahead of the hands. However, the cognitive strategy while sight-reading serial music may involve the immediate conversion of notes being read into finger movements rather than utilizing pattern recognition based on accumulated musical knowledge in long-term memory. This could manifest as a very short EHS, where the hands closely follow the reading eyes. The reason is that in music, where pattern recognition is challenging, increasing the gap between the eyes and the hands can lead to cognitive overload from an information storage perspective.

For the ecological validity of the experiment, quantitatively controlling and ensuring the objectivity of the sight-reading materials is also a crucial aspect of the research. Processing and controlling sight-reading materials with a diverse range of musical styles for experimental use is undeniably a challenging task. However, ultimately, by using a variety of stylistically different pieces that reflect (or are similar to) real sight-reading situations to measure the EHS and investigating how the EHS is extended or shortened based on the structural and stylistic characteristics of the music, it would be possible to better explain the cognitive strategies employed by performers while sight-reading within real-world scenarios.

Aside from musical styles, several studies suggested that the size of the EHS can vary based on multiple musical variables, albeit not musical styles (complexity in Cara, 2018; Rosemann et al., 2016; Wurtz et al., 2009; difficulty in Imai-Matsumura & Mutou, 2021, 2023; structure in Cara, 2023; Penttinen et al., 2015; texture in Weaver, 1943). However, they have not directly explored the interaction effect of the variables (musical domain) on the relationship between the EHS (cognitive domain) and sight-reading proficiency (behavioral domain) and primarily centered on the relationship between the EHS and musical domain within this study's framework. The primary interest of this study is in understanding how the interplay between the EHS and sight-reading proficiency interacts with subfactors within the musical domain. If future research probes into whether or not the interrelation between the EHS and sight-reading proficiency changes with various musical variables, including different styles of music and not just based on task difficulty, and examines how such variations impact the EHS strategy, it can pave the way for a more detailed and comprehensive model to fathom the intricacies of sight-reading proficiency and the underlying mechanisms.

Chapter 6. Conclusion

The dissertation began with a definition of sight-reading and concludes its comprehensive journey with a detailed exploration of a particular facet: the EHS. This scholarly endeavor attempted to answer the research question concerning the reasons for individual differences in sight-reading proficiency among professional performers. This dissertation investigated this research question through the lens of cognitive psychology, involving empirical research through eye tracking. It proposed a conceptual framework, which segmented sight-reading into musical (complexity and playing tempo), cognitive (EHS), and behavioral (performance accuracy) domains and examined the interplay among them.

The findings suggest that the EHS, while significant, is not a decisive indicator of sight-reading proficiency. Rather, it serves as a flexible strategy tailored to varying musical challenges. Proficient sight-readers did not consistently demonstrate a longer EHS. Less-skilled readers did not display a consistently shorter EHS. Instead, proficient sight-readers were observed to continuously adjust their span length in response to the perceived difficulty of the sight-reading tasks. For example, during less challenging tasks, extended EHS, where the eyes look farther ahead than the hands, proved advantageous for proficient sight-reading, corroborating traditional perspectives to some extent. By contrast, as task difficulty increased, successful sight-readers tended to reduce their EHS, thus maintaining closer visual-motor congruence—with the eyes looking precisely note-by-note and closely following the hands. Performers with superior sight-reading abilities demonstrated greater adeptness at leveraging this variability based on the task characteristics. These findings challenge "*conventional wisdom* [emphasis added] about sight-reading among musicians, which [posits] that good sight-readers should have their eyes well ahead of the hands" (Truitt et al., 1997, p. 160).

The observed variability in the EHS implies a dynamic interaction between the cognitive and behavioral domains within the musical context, underscoring the need for a multidimensional approach to sight-reading proficiency. The empirical investigation provides a possible answer to the dissertation's research question: Those proficient in sight-reading flexibly expand and contract their EHS based on musical contexts and challenges with greater agility and precision, highlighting the importance of adaptability as an essential *strategy* in sight-reading.

This dissertation outlines a key avenue for future research in sight-reading proficiency beyond the EHS. There is a compelling need to endeavor further into a neuroscientific approach identifying the specific brain regions associated with sight-reading abilities. Despite a sustained inquiry into the neural mechanism underlying reading music, most studies have focused on silent music reading—reading a score without performing it (Gunter et al., 2003; Hoppe et al., 2014; Nichols & Grahn, 2016; Roux et al., 2007; Schön & Besson, 2003; Simoens & Tervaniemi, 2013; Wong et al., 2014; Wong & Gauthier, 2010)—and have contrasted musicians with nonmusicians across a broad skill spectrum (Behmer & Jantzen, 2011; Hoppe et al., 2014; Lee & Lei, 2012; Li & Hsiao, 2018; Nichols & Grahn, 2016). Whereas a limited number of studies (e.g., Sergent et al., 1992; Schön et al., 2002) have identified brain regions associated with sight-reading *performance*, such as the right occipitotemporal junction, they have not examined how the neural activation in the region(s) varies as a function of sight-reading abilities. Identifying the potential brain areas involved in sight-reading and understanding how these areas respond differently according to varying levels of sight-reading skills can explain the neurological underpinnings of sight-reading proficiency.

For instance, in language research, words are selectively processed within the visual word form area (VWFA) in the left ventral occipitotemporal cortex (VOTC) (Carreiras et al., 2014; Dehaene & Cohen, 2011; Glezer et al., 2009; McCandliss et al., 2003). Studies have shown that reading skill modulates responses in the VOTC; skilled readers display more robust responses to words than objects in this region, indicating increased word selectivity as reading skill improves (Kubota et al., 2019). Interestingly, Schön et al. (2002) suggested that the right occipitotemporal junction, activated prominently while reading musical notations, can be considered a musical analog of the VWFA. This analogy raises the question of whether similar neural activation patterns, for example, higher selectivity of particular indicators such as notes, reflective of sight-reading proficiency, exist within the right occipitotemporal junction. Investigating such a query could constitute a fascinating neuroscientific approach to sight-reading proficiency.

A scientific inquiry into sight-reading offers implications in various aspects. From the cognitive science perspective, empirical research on sight-reading contributes to our broader understanding of advanced reading abilities, elucidating how individuals decode and process complex symbolic information rapidly and accurately. Sight-reading is a multisensory integration that involves translating visual cues into motor responses and monitoring the output through auditory feedback, demonstrating how various sensorimotor modalities coalesce into a cohesive cognitive action. This multimodal nature of sight-reading highlights the simultaneous engagement of advanced cognitive functions such as attention, memory, and motor planning, offering profound insights into the execution of complex cognitive tasks. In terms of cognitive flexibility, sight-reading, which requires the ability to adapt to unfamiliar pieces of music in real time, sheds light on the strategic thinking and cognitive adaptability necessary for managing cognitive load during task performance involving new and changing conditions. From an educational standpoint, exploring the cognitive mechanisms of sight-reading can help music educators devise evidence-based teaching methods and educational environments rather than solely relying on intuition. For instance, studies have shown that different cognitive strategies and abilities may be used based on who sight-reads and what is sight-read (Perra et al., 2021; Puurtinen, 2018). With this knowledge, music educators can analyze sight-reading abilities individually and develop optimized training programs for various skill levels and musical contexts. Finally, from a musical perspective, empirical research on sight-reading is meaningful as it sheds light on the specific abilities that constitute this historically mystified musical ability. Moreover, understanding the cognitive mechanisms of sight-reading proficiency can help alleviate performers' intuitive fear of sight-reading and strategically refine their sight-reading skills.

Bibliography

Ahken, S., Comeau, G., Hébert, S., & Balasubramaniam, R. (2012). Eye movement patterns during the processing of musical and linguistic syntactic incongruities. *Psychomusicology: Music, Mind, and Brain*, *22*(1), 18–25. https://doi.org/10.1037/a0026751

Aiba, E., & Matsui, T. (2016). Music memory following short-term practice and its relationship with the sight-reading abilities of professional pianists. *Frontiers in Psychology*, *7*, 645. https://doi.org/10.3389/fpsyg.2016.00645

Alexander, M. L., & Henry, M. L. (2012). The development of a string sight-reading pitch skill hierarchy. *Journal of Research in Music Education*, *60*(2), 201–216. https://doi.org/10.1177/0022429412446375

Arthur, P., Khuu, S., & Blom, D. (2016). Music sight-reading expertise, visually disrupted score and eye movements. *Journal of Eye Movement Research*, *9*(7), 1–11. https://doi.org/doi:10.16910/jemr.9.7.1

Arthur, P., Khuu, S., & Blom, D. (2021). Visual processing abilities associated with piano music sight-reading expertise. *Psychology of Music*, *49*(4), 1006–1016. https://doi.org/10.1177/0305735620920370

Arthur, P., McPhee, E., & Blom, D. (2020). Determining what expert piano sight-readers have in common. *Music Education Research*, *22*(4), 447–456. https://doi.org/10.1080/14613808.2020.1767559

Ashby, J., Rayner, K., & Clifton Jr, C. (2005). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *The Quarterly Journal of Experimental Psychology Section A*, *58*(6), 1065–1086. https://doi.org/10.1080/02724980443000476

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556–559. https://doi.org/10.1126/science.1736359

Bangert, D., Schubert, E., & Fabian, D. (2009). Decision-making in unpracticed and practiced performances of Baroque violin music. In C. Stevens, E. Schubert, B. Kruithof, K. Buckley & S. Fazio (Eds.), *Proceedings of the International Conference on Music Communication Science* (pp. 3–6). Sydney, Australia: HCSNet, University of Western Sydney.

Banton, L. J. (1995). The role of visual and auditory feedback during the sight-reading of music. *Psychology of Music*, *23*(1), 3–16. https://doi.org/10.1177/0305735695231001

Barenboim, D. (2003). *A life in music*. Arcade Publishing.

Bean, K. L. (1938). An experimental approach to the reading of music. *Psychological Monographs*, *50*(6), i–80. https://doi.org/10.1037/h0093540

Behmer Jr, L. P., & Jantzen, K. J. (2011). Reading sheet music facilitates sensorimotor mu-desynchronization in musicians. *Clinical Neurophysiology*, *122*(7), 1342–1347. https://doi.org/10.1016/j.clinph.2010.12.035

Bélanger, N. N., Slattery, T. J., Mayberry, R. I., & Rayner, K. (2012). Skilled deaf readers have an enhanced perceptual span in reading. *Psychological Science*, *23*(7), 816–823. https://doi.org/10.1177/0956797611435130

Betts, S. L., & Cassidy, J. W. (2000). Development of harmonization and sight-reading skills among university class piano students. *Journal of Research in Music Education*, *48*(2), 151–161. https://doi.org/10.2307/3345573

Bharucha, J. J. (1987). Music cognition and perceptual facilitation: A connectionist framework. *Music Perception*, *5*(1), 1–30. https://doi.org/10.2307/40285384

Bor, D., Duncan, J., Wiseman, R. J., & Owen, A. M. (2003). Encoding strategies dissociate prefrontal activity from working memory demand. *Neuron*, *37*(2), 361–367. https://doi.org/10.1016/S0896-6273(02)01171-6

Boyle, J. D. (1970). The effect of prescribed rhythmical movements on the ability to read music at sight. *Journal of Research in Music Education*, *18*(4), 307–318. https://doi.org/10.2307/3344498

Bozkurt, B., Baysal, O., & Yüret, D. (2017). A dataset and baseline system for singing voice assessment. In *Proceedings of the International Symposium on Computer Music Multidisciplinary Research* (pp. 430–438), Matosinhos, Portugal.

Buck, P. C. (1944). *Psychology for musicians*. Oxford University Press.

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

Buswell, G. T. (1920). *An experimental study of the eye-voice span in reading*. University of Chicago Press.

Butsch, R. L. (1932). Eye movements and the eye-hand span in typewriting. *Journal of Educational Psychology*, *23*(2), 104–121. https://doi.org/10.1037/h0073463

Cara, M. A. (2023). The effect of practice and musical structure on pianists' eye-hand span and visual monitoring. *Journal of Eye Movement Research*, *16*(2). https://doi.org/10.16910/jemr.16.2.5

Cara, M. A. (2018). Anticipation awareness and visual monitoring in reading contemporary music. *Musicae Scientiae*, *22*(3), 322–343. https://doi.org/10.1177/1029864916687601

Cara, M. A., & Gómez, G. (2016). Silent reading of music and texts; eye movements and integrative reading mechanisms. *Journal of Eye Movement Research*, *9*(7). https://doi.org/10.16910/jemr.9.7.2

Carreiras, M., Armstrong, B. C., Perea, M., & Frost, R. (2014). The what, when, where, and how of visual word recognition. *Trends in Cognitive Sciences*, *18*(2), 90–98. https://doi.org/10.1016/j.tics.2013.11.005

Chaffin, R., & Imreh, G. (2002). Practicing perfection: Piano performance as expert memory. *Psychological Science*, *13*(4), 342–349. https://doi.org/10.1111/j.0956-7976.2002.00462.x

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*(1), 55–81. https://doi.org/10.1016/0010-0285(73)90004-2

Chitalkina, N., Puurtinen, M., Gruber, H., & Bednarik, R. (2021). Handling of incongruences in music notation during singing or playing. *International Journal of Music Education*, *39*(1), 18–38. https://doi.org/10.1177/0255761420944036

Choi, W., Lowder, M. W., Ferreira, F., & Henderson, J. M. (2015). Individual differences in the perceptual span during reading: Evidence from the moving window technique. *Attention, Perception, & Psychophysics*, *77*, 2463–2475. https://doi.org/10.3758/s13414-015-0942-1

Chun, H. (2022). Exploring effective sight-reading methods for intermediate-level pianists in auditory, visual and technical proficient aspects: A literature review. *Frontiers in Art Research*, *4*(12), 94–98. https://doi.org/10.25236/FAR.2022.041219

Clarke, E. F. (1987). Levels of structure in the organization of musical time. *Contemporary Music Review*, *2*(1), 211–238. https://doi.org/10.1080/07494468708567059

Colwell, R. J. (1969). *The teaching of instrumental music*. Prentice-Hall.

Conklin, D., & Witten, I. H. (1995). Multiple viewpoint systems for music prediction. *Journal of New Music Research*, *24*(1), 51–73. https://doi.org/10.1080/09298219508570672

Cook, N. (2009). Methods for analysing recordings. In N. Cook, E. Clarke, D. Leech-Wilkinson, & J. Rink (Eds.), *The Cambridge companion to recorded music* (pp. 221–245). Cambridge University Press. https://doi.org/10.1017/CCOL9780521865821.027

Cox, B. E. (2000). *Factors associated with success in sight reading four -part chordal piano music* (Order No. 9958982) [Doctoral dissertation, Auburn University]. ProQuest Dissertations and Theses Global.

D'Anselmo, A., Giuliani, F., Marzoli, D., Tommasi, L., & Brancucci, A. (2015). Perceptual and motor laterality effects in pianists during music sight-reading. *Neuropsychologia*, *71*, 119–125. https://doi.org/10.1016/j.neuropsychologia.2015.03.026

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, *15*(6), 254–262. https://doi.org/10.1016/j.tics.2011.04.003

Delogu, F., Brunetti, R., Inuggi, A., Campus, C., Del Gatto, C., & D'Ausilio, A. (2019). That does not sound right: Sounds affect visual ERPs during a piano sight-reading task. *Behavioural Brain Research*, *367*, 1–9. https://doi.org/10.1016/j.bbr.2019.03.037

Dib, N. E., & Sturmey, P. (2011). Effects of general-case training, instructions, rehearsal, and feedback on the reduction of sight-reading errors by competent musicians. *Journal of Applied Behavior Analysis*, *44*(3), 599–604. https://doi.org/10.1901/jaba.2011.44-599

Drai-Zerbib, V., & Baccino, T. (2014). The effect of expertise in music reading: Cross-modal competence. *Journal of Eye Movement Research*, *6*(5), 1–10.

Drai-Zerbib, V., Baccino, T., & Bigand, E. (2012). Sight-reading expertise: Cross-modality integration investigated using eye tracking. *Psychology of Music*, *40*(2), 216–235. https://doi.org/10.1177/0305735610394710

Drake, C., & Palmer, C. (2000). Skill acquisition in music performance: Relations between planning and temporal control. *Cognition*, *74*(1), 1–32. https://doi.org/10.1016/S0010-0277(99)00061-X

Eaton, J. L. (1978). *A correlation study of keyboard sight-reading facility with previous training, note-reading, psychomotor, and memorization skills* (Order No. 7901575) [Doctoral dissertation, Indiana University]. ProQuest Dissertations and Theses Global.

Elliott, C. A. (1982). The relationships among instrumental sight-reading ability and seven selected predictor variables. *Journal of Research in Music Education*, *30*(1), 5–14. https://doi.org/10.2307/3344862

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*(1), 19–23. https://doi.org/10.1111/1467-8721.00160

Ericsson, K. A. (1996). The acquisition of expert performance: An introduction to some of the issues. In K. A. Ericsson (Ed.), *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games* (pp. 1–50). Lawrence Erlbaum Associates, Inc.

Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist*, *49*(8), 725–747. https://doi.org/10.1037/0003-066X.49.8.725

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*(2), 211–245. https://doi.org/10.1037/0033-295X.102.2.211

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363–406. https://doi.org/10.1037/0033-295X.100.3.363

Fabian, D., Timmers, R., & Schubert, E. (Eds.). (2014). *Expressiveness in music performance: Empirical approaches across styles and cultures*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199659647.001.0001

Fan, P., Wong, A. C. N., & Wong, Y. K. (2022). Visual and visual association abilities predict skilled reading performance: The case of music sight-reading. *Journal of Experimental Psychology: General*, *151*(11), 2683–2705. https://doi.org/10.1037/xge0001217

Fine, P., Berry, A., & Rosner, B. (2006). The effect of pattern recognition and tonal predictability on sight-singing ability. *Psychology of Music*, *34*(4), 431–447. https://doi.org/10.1177/0305735606067152

Furneaux, S., & Land, M. F. (1999). The effects of skill on the eye–hand span during musical sight–reading. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *266*(1436), 2435–2440. https://doi.org/10.1098/rspb.1999.0943

Gilman, E., & Underwood, G. (2003). Restricting the field of view to investigate the perceptual spans of pianists. *Visual Cognition*, *10*(2), 201–232. https://doi.org/10.1080/713756679

Glezer, L. S., Jiang, X., & Riesenhuber, M. (2009). Evidence for highly selective neuronal tuning to whole words in the "visual word form area". *Neuron*, *62*(2), 199–204. https://doi.org/10.1016/j.neuron.2009.03.017

Gobet, F. (1998). Memory for the meaningless: How chunks help. In M. A. Gernsbacker & S. J. Derry (Eds.), *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (pp. 398–403). Lawrence Erlbaum Associates.

Goolsby, T. W. (1994a). Eye movement in music reading: Effects of reading ability, notational complexity, and encounters. *Music Perception*, *12*(1), 77–96. https://doi.org/10.2307/40285756

Goolsby, T. W. (1994b). Profiles of processing: Eye movements during sightreading. *Music Perception*, *12*(1), 97–123. https://doi.org/10.2307/40285757

Gromko, J. E. (2004). Predictors of music sight-reading ability in high school wind players. *Journal of Research in Music Education*, *52*(1), 6–15. https://doi.org/10.2307/3345521

Grutzmacher, P. A. (1987). The effect of tonal pattern training on the aural perception, reading recognition, and melodic sight-reading achievement of first-year instrumental music students. *Journal of Research in Music Education*, *35*(3), 171–181. https://doi.org/10.2307/3344959

Gunter, T. C., Schmidt, B. H., & Besson, M. (2003). Let's face the music: A behavioral and electrophysiological exploration of score reading. *Psychophysiology*, *40*(5), 742–751. https://doi.org/10.1111/1469-8986.00074

Hadley, L. V., Sturt, P., Eerola, T., & Pickering, M. J. (2018). Incremental comprehension of pitch relationships in written music: Evidence from eye movements. *The Quarterly Journal of Experimental Psychology*, *71*(1), 211–219. https://doi.org/10.1080/17470218.2017.1307861

Häikiö, T., Bertram, R., Hyönä, J., & Niemi, P. (2009). Development of the letter identity span in reading: Evidence from the eye movement moving window paradigm. *Journal of Experimental Child Psychology*, *102*(2), 167–181. https://doi.org/10.1016/j.jecp.2008.04.002

Hayward, C. M., & Eastlund Gromko, J. (2009). Relationships among music sight-reading and technical proficiency, spatial visualization, and aural discrimination. *Journal of Research in Music Education*, *57*(1), 26–36. https://doi.org/10.1177/0022429409332677

Herrero, L., & Carriedo, N. (2019). The contributions of updating in working memory sub-processes for sight-reading music beyond age and practice effects. *Frontiers in Psychology*, *10*, 1080. https://doi.org/10.3389/fpsyg.2019.01080

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science*, *143*(3611), 1190–1192. https://doi.org/10.1126/science.143.3611.1190

Hillbrand, E. K. (1924). *Measuring ability in sight singing* (Unpublished doctoral dissertation). Northewestern University.

Hiller, L., & Fuller, R. (1967). Structure and information in Webern's Symphonie, Op. 21. *Journal of Music Theory*, *11*(1), 60–115. https://doi.org/10.2307/842949

Hoppe, C., Splittstößer, C., Fliessbach, K., Trautner, P., Elger, C. E., & Weber, B. (2014). Silent music reading: Auditory imagery and visuotonal modality transfer in singers and non-singers. *Brain and Cognition*, *91*, 35–44. https://doi.org/10.1016/j.bandc.2014.08.002

Huovinen, E., Ylitalo, A. K., & Puurtinen, M. (2018). Early attraction in temporally controlled sight reading of music. *Journal of Eye Movement Research*, *11*(2), 10.16910/jemr.11.2.3. https://doi.org/10.16910/jemr.11.2.3

Imai-Matsumura, K., & Mutou, M. (2021). Gaze analysis of pianists' sight-reading: comparison between expert pianists and students training to be pianists. *Music & Science*, *4*. https://doi.org/10.1177/20592043211061106

Imai-Matsumura, K., & Mutou, M. (2023). The influence of executive functions on eye-hand span and piano performance during sight-reading. *PLOS ONE*, *18*(5), e0285043. https://doi.org/10.1371/journal.pone.0285043

Inhoff, A. W., Solomon, M., Radach, R., & Seymour, B. A. (2011). Temporal dynamics of the eye–voice span and eye movement control during oral reading. *Journal of Cognitive Psychology*, *23*(5), 543–558. https://doi.org/10.1080/20445911.2011.546782

Inhoff, A. W., Starr, M., & Shindler, K. L. (2000). Is the processing of words during eye fixations in reading strictly serial?. *Perception & Psychophysics*, *62*(7), 1474–1484. https://doi.org/10.3758/BF03212147

Iorio, C., Šaban, I., Poulin-Charronnat, B., & Schmidt, J. R. (2023). Incidental learning in music reading: The music contingency learning task. *The Quarterly Journal of Experimental Psychology*, *76*(2), 429–449. https://doi.org/10.1177/17470218221092779

Jacobs, A. M. (1986). Eye-movement control in visual search: How direct is visual span control?. *Perception & Psychophysics*, *39*(1), 47–58. https://doi.org/10.3758/BF03207583

Jacobsen, O. I. (1928). An experimental study of photographing eye-movements in reading music. *Music Supervisors' Journal*, *14*(3), 63–69. https://doi.org/10.2307/3382999

Kim, Y. J., Song, M. K., & Atkins, R. (2021). What is your thought process during sight-reading? Advanced sight-readers' strategies across different tonal environments. *Psychology of Music*, *49*(5), 1303–1320. https://doi.org/10.1177/0305735620942596

Kinsler, V., & Carpenter, R. H. S. (1995). Saccadic eye movements while reading music. *Vision Research*, *35*(10), 1447–1458. https://doi.org/10.1016/0042-6989(95)98724-N

Knopoff, L., & Hutchinson, W. (1983). Entropy as a measure of style: The influence of sample length. *Journal of Music Theory*, *27*(1), 75–97. https://doi.org/10.2307/843561

Koelsch, S., Rohrmeier, M., Torrecuso, R., & Jentschke, S. (2013). Processing of hierarchical syntactic structure in music. *Proceedings of the National Academy of Sciences*, *110*(38), 15443–15448. https://doi.org/10.1073/pnas.1300272110

Kopiez, R., & In Lee, J. (2006). Towards a dynamic model of skills involved in sight reading music. *Music Education Research*, *8*(1), 97–120. https://doi.org/10.1080/14613800600570785

Kopiez, R., & In Lee, J. (2008). Towards a general model of skills involved in sight reading music. *Music Education Research*, *10*(1), 41–62. https://doi.org/10.1080/14613800701871363

Kopiez, R., Weihs, C., Ligges, U., & Lee, J. I. (2006). Classification of high and low achievers in a music sight-reading task. *Psychology of Music*, *34*(1), 5–26. https://doi.org/10.1177/0305735606059102

Kornicke, L. E. (1992). *An exploratory study of individual difference variables in piano sight-reading achievement* (Order No. 9301458) [Doctoral dissertation, Indiana University]. ProQuest Dissertations and Theses Global.

Kostka, M. J. (2000). The effects of error-detection practice on keyboard sight-reading achievement of undergraduate music majors. *Journal of Research in Music Education*, *48*(2), 114–122. https://doi.org/10.2307/3345570

Krejtz, K., Duchowski, A. T., Niedzielska, A., Biele, C., & Krejtz, I. (2018). Eye tracking cognitive load using pupil diameter and microsaccades with fixed gaze. *PLOS ONE*, *13*(9), e0203629. https://doi.org/10.1371/journal.pone.0203629

Krieber, M., Bartl-Pokorny, K. D., Pokorny, F. B., Einspieler, C., Langmann, A., Körner, C., ... & Marschik, P. B. (2016). The relation between reading skills and eye movement patterns in adolescent readers: Evidence from a regular orthography. *PLOS ONE*, *11*(1), e0145934. https://doi.org/10.1371/journal.pone.0145934

Krumhansl, C. L. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin*, *126*(1), 159–179. https://doi.org/10.1037/0033-2909.126.1.159

Kubota, E. C., Joo, S. J., Huber, E., & Yeatman, J. D. (2019). Word selectivity in high-level visual cortex and reading skill. *Developmental Cognitive Neuroscience*, *36*, 100593. https://doi.org/10.1016/j.dcn.2018.09.003

Kuperman, V., & Van Dyke, J. A. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language*, *65*(1), 42–73. https://doi.org/10.1016/j.jml.2011.03.002

Lannert, V., & Ullman, M. (1945). Factors in the reading of piano music. *The American Journal of Psychology*, *58*(1), 91–99. https://doi.org/10.2307/1417577

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior; the Hixon Symposium* (pp. 112–146). Wiley.

Lawson, E. A. (1961). A note on the influence of different orders of approximation to the English language upon eye-voice span. *The Quarterly Journal of Experimental Psychology*, *13*(1), 53–55. https://doi.org/10.1080/17470216108416469

Levin, H., & Kaplan, E. A. (1970). Grammatical structure and reading. In H. Levin & J. P. Williams (Eds.), *Basic studies on reading* (pp. 119–133). Basic Books.

Lee, H. Y., & Lei, S. F. (2012). Musical training effect on reading musical notation: evidence from event-related potentials. *Perceptual and Motor Skills*, *115*(1), 7–17. https://doi.org/10.2466/22.11.24.PMS.115.4.7-17

Lehmann, A. C., & Ericsson, K. A. (1993). Sight-reading ability of expert pianists in the context of piano accompanying. *Psychomusicology: A Journal of Research in Music Cognition*, *12*(2), 182–195. https://doi.org/10.1037/h0094108

Lehmann, A. C., & Ericsson, K. A. (1996). Performance without preparation: Structure and acquisition of expert sight-reading and accompanying performance. *Psychomusicology: A Journal of Research in Music Cognition*, *15*(1–2), 1–29. https://doi.org/10.1037/h0094082

Lehmann, A. C., & Kopiez, R. (2016). Sight-reading. In S. Hallam, I. Cross, & M. H. Thaut (Eds.), *The Oxford handbook of music psychology* (2nd ed., pp. 547–558). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198722946.013.33

Lehmann, A. C., & McArthur, V. (2002). Sight-reading. In R. Parncutt & Gary McPherson (Eds.), *The science and psychology of music performance: Creative strategies for teaching and learning* (pp. 135–150). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195138108.003.0009

Lehmann, A. C., Sloboda, J. A., Woody, R. H. (2007). Reading or listening and remebering. In A. C. Lehmann, J. A. Sloboda, & R. H. Woody (Eds.), *Psychology for musicians: Understanding and acquiring the skills* (pp. 107–126). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195146103.003.0006

Li, S. T. K., & Hsiao, J. H. W. (2018). Music reading expertise modulates hemispheric lateralization in English word processing but not in Chinese character processing. *Cognition*, *176*, 159–173. https://doi.org/10.1016/j.cognition.2018.03.010

Lim, Y., Park, J. M., Rhyu, S. Y., Chung, C. K., Kim, Y., & Yi, S. W. (2019). Eye-hand span is not an indicator of but a strategy for proficient sight-reading in piano performance. *Scientific Reports*, *9*(1), 17906. https://doi.org/10.1038/s41598-019-54364-y

Lörch, L. (2021). The association of eye movements and performance accuracy in a novel sight-reading task. *Journal of Eye Movement Research*, *14*(4), 10.16910/jemr.14.4.5. https://doi.org/10.16910/jemr.14.4.5

Lowder, J. E. (1983). Evaluation of keyboard skills required in college class piano programs. *Contributions to Music Education*, 10, 33–38. http://www.jstor.org/stable/24127376

Lu, C.-I., Greenwald, M. L., Lin, Y.-Y., & Bowyer, S. M. (2021). Musical transposing versus sight-reading: Mapping brain activation with magnetoencephalography. *Psychology of Music*, *49*(3), 581–599. https://doi.org/10.1177/0305735619883692

Luce, J. R. (1965). *Sight-reading and ear-playing abilities as related to instrumental music students*. *Journal of Research in Music Education*, *13*(2), 101–109. https://doi.org/10.2307/3344447

Lyke, J. B. (1968). *An investigation of class piano programs in the six state universities of Illinois and recommendations for their improvement* (Publication No. AAT 6814732) [Doctoral dissertation, University of Nothern Colorado]. Dissertation Abstracts International.

MacGregor, J. N. (1987). Short-term memory capacity: Limitation or optimization?. *Psychological Review*, *94*(1), 107–108. https://doi.org/10.1037/0033-295X.94.1.107

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

Maturi, K. S., & Sheridan, H. (2020). Expertise effects on attention and eye-movement control during visual search: Evidence from the domain of music reading. *Attention, Perception, & Psychophysics*, *82*, 2201–2208. https://doi.org/10.3758/s13414-020-01979-3

McAdams, S. (1989). Psychological constraints on form-bearing dimensions in music. *Contemporary Music Review*, *4*(1), 181–198. https://doi.org/10.1080/07494468900640281

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences*, *7*(7), 293–299. https://doi.org/10.1016/S1364-6613(03)00134-7

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics*, *17*, 578–586. https://doi.org/10.3758/BF03203972

McPherson, G. E. (1994). Factors and abilities influencing sightreading skill in music. *Journal of Research in Music Education*, *42*(3), 217–231. https://doi.org/10.2307/3345701

Meinz, E. J., & Hambrick, D. Z. (2010). Deliberate practice is necessary but not sufficient to explain individual differences in piano sight-reading skill: The role of working memory capacity. *Psychological Science*, *21*(7), 914–919. https://doi.org/10.1177/0956797610373933

Miller, R. E. (1988). *Contributions of selected music skills to music sight-reading achievement and rehearsed reading achievement* (Unpublished doctoral dissertation). University of Illinois at Urbana-Champaign.

Mishra, J. (2014a). Improving sightreading accuracy: A meta-analysis. *Psychology of Music*, *42*(2), 131–156. https://doi.org/10.1177/0305735612463770

Mishra, J. (2014b). Factors related to sight-reading accuracy: A meta-analysis. *Journal of Research in Music Education*, *61*(4), 452–465. https://doi.org/10.1177/0022429413508585

Mishra, J. (2016). Rhythmic and melodic sight reading interventions: Two meta-analyses. *Psychology of Music*, *44*(5), 1082–1094. https://doi.org/10.1177/0305735615610925

Molina, E., Barbancho, I., Gómez, E., Barbancho, A. M., & Tardón, L. J. (2013). Fundamental frequency alignment vs. note-based melodic similarity for singing voice assessment. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 744–748), Vancouver, Canada. https://doi.org/10.1109/ICASSP.2013.6637747

Morton, J. (1964). The effects of context upon speed of reading, eye movements and eye-voice span. *The Quarterly Journal of Experimental Psychology*, *16*(4), 340–354. https://doi.org/10.1080/17470216408416390

Moussard, A., Bigand, E., Belleville, S., & Peretz, I. (2012). Music as an aid to learn new verbal information in Alzheimer's disease. *Music Perception*, *29*(5), 521–531. https://doi.org/10.1525/mp.2012.29.5.521

Müller, M. (2007). Dynamic time warping. In M. Müller (Ed.), *Information retrieval for music and motion* (pp. 69–84). Springer. https://doi.org/10.1007/978-3-540-74048-3_4

Nakamura, E., Yoshii, K., & Katayose, H. (2017). Performance error detection and post-processing for fast and accurate symbolic music alignment. In *Proceedings of the 18th International Conference on Music Information Retrieval* (pp. 347–353), Suzhou, China.

Nichols, E. S., & Grahn, J. A. (2016). Neural correlates of audiovisual integration in music reading. *Neuropsychologia*, *91*, 199–210. https://doi.org/10.1016/j.neuropsychologia.2016.08.011

Norris, D., Kalm, K., & Hall, J. (2020). Chunking and redintegration in verbal short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(5), 872–893. https://doi.org/10.1037/xlm0000762

Olsen A. (2012). Tobii I-VT fixation filter—Algorithm description. *White Paper*. Retrieved from http://www.tobii.com/eye-tracking-research/global/library/white-papers/the-tobii-i-vt-fixation-filter/

Ortmann, O. (1934). Elements of chord-reading in music notation. *The Journal of Experimental Education*, *3*(1), 50–57. https://doi.org/10.1080/00220973.1934.11009965

Pan, J., Li, M., Song, Z., Li, X., Liu, X., Yi, H., & Zhu, M. (2017). An audio based piano performance evaluation method using deep neural network based acoustic modeling. In *Proceedings of Interspeech 2017* (pp. 3088–3092). https://doi.org/10.21437/Interspeech.2017-866

Penttinen, M. (2013). *Skill development in music reading. The eye-movement approach* [Doctoral dissertation, University of Turku]. https://www.utupub.fi/bitstream/handle/10024/88757/AnnalesB359PenttinenDISS.pdf?sequence=1

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

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

Penttinen, M., Huovinen, E., & Ylitalo, A. K. (2015). Reading ahead: Adult music students' eye movements in temporally controlled performances of a children's song. *International Journal of Music Education*, *33*(1), 36–50. https://doi.org/10.1177/0255761413515813

Perra, J., Latimier, A., Poulin-Charronnat, B., Baccino, T., & Drai-Zerbib, V. (2022). A meta-analysis on the effect of expertise on eye movements during music reading. *Journal of Eye Movement Research*, *15*(4), 10.16910/jemr.15.4.1. https://doi.org/10.16910/jemr.15.4.1

Perra, J., Poulin-Charronnat, B., Baccino, T., & Drai-Zerbib, V. (2021). Review on eye-hand span in sight-reading of music. *Journal of Eye Movement Research*, *14*(4), 10.16910/jemr.14.4.4. https://doi.org/10.16910/jemr.14.4.4

Pike, P. D., & Carter, R. (2010). Employing cognitive chunking techniques to enhance sight-reading performance of undergraduate group-piano students. *International Journal of Music Education*, *28*(3), 231–246. https://doi.org/10.1177/0255761410373886

Platz, F., Kopiez, R., Lehmann, A. C., & Wolf, A. (2014). The influence of deliberate practice on musical achievement: a meta-analysis. *Frontiers in Psychology*, *5*, 646. https://doi.org/10.3389/fpsyg.2014.00646

Polanka, M. (1995). Research note: Factors affecting eye movements during the reading of short melodies. *Psychology of Music*, *23*(2), 177–183. https://doi.org/10.1177/0305735695232005

Pomerleau-Turcotte, J., Dubé, F., Moreno Sala, M. T., & Vachon, F. (2023). Building a mental toolbox: Relationships between strategy choice and sight-singing performance in higher education. *Psychology of Music*, *51*(1), 119–139. https://doi.org/10.1177/03057356221087444

Puurtinen, M. (2018). Eye on music reading: A methodological review of studies from 1994 to 2017. *Journal of Eye Movement Research*, *11*(2), 10.16910/jemr.11.2.2. https://doi.org/10.16910/jemr.11.2.2

Puurtinen, M., Huovinen, E., & Ylitalo, A. K. (2023). Cognitive mechanisms in temporally controlled rhythm reading: Evidence from eye movements. *Music Perception*, *40*(3), 237–252. https://doi.org/10.1525/mp.2023.40.3.237

Qi, J., & Adachi, M. (2022). The influence of modality on input, visuo-motor coordination, and execution in the advanced pianist's sight-reading processes. *Frontiers in Psychology*, *13*, 933106. https://doi.org/10.3389/fpsyg.2022.933106

Quantz, J. O. (1897). Problems in the psychology of reading. *The Psychological Review: Monograph Supplements*, *2*(1), i–51. https://doi.org/10.1037/h0092985

Rabinovich, M. I., Varona, P., Tristan, I., & Afraimovich, V. S. (2014). Chunking dynamics: Heteroclinics in mind. *Frontiers in Computational Neuroscience*, *8*, 22. https://doi.org/10.3389/fncom.2014.00022

Radach, R., & Kennedy, A. (2004). Theoretical perspectives on eye movements in reading: Past controversies, current issues, and an agenda for future research. *European Journal of Cognitive Psychology*, *16*(1–2), 3–26. https://doi.org/10.1080/09541440340000295

Radach, R., & Kennedy, A. (2013). Eye movements in reading: Some theoretical context. *The Quarterly Journal of Experimental Psychology*, *66*(3), 429–452. https://doi.org/10.1080/17470218.2012.750676

Randel, D. M. (2003). *The Harvard dictionary of* music (4th ed.). Belknap Press of Harvard University Press.

Raney, G. E., Campbell, S. J., & Bovee, J. C. (2014). Using eye movements to evaluate the cognitive processes involved in text comprehension. *Journal of Visualized Experiments*, *83*, e50780. https://doi.org/10.3791/50780

Rayner, K. (1977). Visual attention in reading: Eye movements reflect cognitive processes. *Memory & Cognition*, *5*(4), 443–448. https://doi.org/10.3758/BF03197383

Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology*, *41*(2), 211–236. https://doi.org/10.1016/0022-0965(86)90037-8

Rayner, K. (1997). Understanding eye movements in reading. *Scientific Studies of Reading*, *1*(4), 317–339. https://doi.org/10.1207/s1532799xssr0104_2

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422. https://doi.org/10.1037/0033-2909.124.3.372

Rayner, K. (1999). What have we learned about eye movements during reading. In R. M. Klein & P. A. McMullen (Eds.), *Converging methods for understanding reading and dyslexia* (pp. 23–56). The MIT Press.

Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, *62*(8), 1457–1506. https://doi.org/10.1080/17470210902816461

Rayner, K., & McConkie, G. W. (1976). What guides a reader's eye movements?. *Vision Research*, *16*(8), 829–837. https://doi.org/10.1016/0042-6989(76)90143-7

Rayner, K., & Pollatsek, A. (1997). Eye movements, the eye-hand span, and the perceptual span during sight-reading of music. *Current Directions in Psychological Science*, *6*(2), 49–53. https://doi.org/10.1111/1467-8721.ep11512647

Rayner, K., & Pollatsek, A. (2006). Eye-movement control in reading. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (pp. 613–658). Academic Press. https://doi.org/10.1016/B978-012369374-7/50017-1

Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific Studies of Reading*, *10*(3), 241–255. https://doi.org/10.1207/s1532799xssr1003_3

Rayner, K., Inhoff, A. W., Morrison, R. E., Slowiaczek, M. L., & Bertera, J. H. (1981). Masking of foveal and parafoveal vision during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(1), 167–179. https://doi.org/10.1037/0096-1523.7.1.167

Rayner, K., Murphy, L. A., Henderson, J. M., & Pollatsek, A. (1989). Selective attentional dyslexia. *Cognitive Neuropsychology*, *6*(4), 357–378. https://doi.org/10.1080/02643298908253288

Rayner, K., Slattery, T. J., & Bélanger, N. N. (2010). Eye movements, the perceptual span, and reading speed. *Psychonomic Bulletin & Review*, *17*(6), 834–839. https://doi.org/10.3758/PBR.17.6.834

Rayner, K., Yang, J., Schuett, S., & Slattery, T. J. (2013). Eye movements of older and younger readers when reading unspaced text. *Experimental Psychology*, *60*(5), 354–361. https://doi.org/10.1027/1618-3169/a000207

Reichle, E. D., Liversedge, S. P., Drieghe, D., Blythe, H. I., Joseph, H. S., White, S. J., & Rayner, K. (2013). Using EZ Reader to examine the concurrent development of eye-movement control and reading skill. *Developmental Review*, *33*(2), 110–149. https://doi.org/10.1016/j.dr.2013.03.001

Repp, B. H. (1990). Patterns of expressive timing in performances of a Beethoven minuet by nineteen famous pianists. *Journal of the Acoustical Society of America*, *88*(2), 622–641. https://doi.org/10.1121/1.399766

Repp, B. H. (1998). A microcosm of musical expression: I. Quantitative analysis of pianists' timing in the initial measures of Chopin's Etude in E major. *Journal of the Acoustical Society of America*, *104*(2, Pt 1), 1085–1100. https://doi.org/10.1121/1.423325

Repp, B. H. (1999). A microcosm of musical expression: II. Quantitative analysis of pianists' dynamics in the initial measures of Chopin's Etude in E major. *Journal of the Acoustical Society of America*, *105*(3), 1972–1988. https://doi.org/10.1121/1.426743

Rink, J. (1990). [Review of the book Musical Structure and Performance, by W. Berry]. *Music Analysis*, *9*(3), 319–339. https://doi.org/10.2307/853982

Rink, J. (2002). Analysis and (or?) performance. In J. Rink (Ed.), *Musical performance: A guide to understanding* (pp. 35–58). Cambridge University Press. https://doi.org/10.1017/CBO9780511811739.004

Rosemann, S., Altenmüller, E., & Fahle, M. (2016). The art of sight-reading: Influence of practice, playing tempo, complexity and cognitive skills on the eye–hand span in pianists. *Psychology of Music*, *44*(4), 658–673. https://doi.org/10.1177/0305735615585398

Roux, F. E., Lubrano, V., Lotterie, J. A., Giussani, C., Pierroux, C., & Demonet, J. F. (2007). When “Abegg” is read and (“A, B, E, G, G”) is not: A cortical stimulation study of musical score reading. *Journal of Neurosurgery*, *106*(6), 1017–1027. https://doi.org/10.3171/jns.2007.106.6.1017

Russell, C. R. (2019). Effects of pitch and rhythm priming tasks on accuracy and fluency during sight-reading. *Journal of Research in Music Education*, *67*(3), 252–269. https://doi.org/10.1177/0022429419851112

Salis, D. L. (1977). *The identification and assessment of cognitive variables associated with reading of advanced music at the piano* (Unpusbliehd doctoral dissertation). University of Pittsburgh.

Salzberg, R. S., & Wang, C. C. (1989). A comparison of prompts to aid rhythmic sight-reading of string students. *Psychology of Music*, *17*(2), 123–131. https://doi.org/10.1177/0305735689172003

Schleuter, S. L. (1993). The relationship of AMMA scores to sightsinging, dictation, and SAT scores of university music majors. *Contributions to Music Education*, 57–63. https://www.jstor.org/stable/24127331

Schön, D., & Besson, M. (2002). Processing pitch and duration in music reading: A RT–ERP study. *Neuropsychologia*, *40*(7), 868–878. https://doi.org/10.1016/s0028-3932(01)00170-1

Schön, D., & Besson, M. (2003). Audiovisual interactions in music reading: A reaction times and event-related potentials study. *Annals of the New York Academy of Sciences*, *999*(1), 193–198. https://doi.org/10.1196/annals.1284.028

Schön, D., Anton, J. L., Roth, M., & Besson, M. (2002). An fMRI study of music sight-reading. *Neuroreport*, *13*(17), 2285–2289. https://doi.org/10.1097/00001756-200212030-00023

Sergent, J., Zuck, E., Terriah, S., & MacDonald, B. (1992). Distributed neural network underlying musical sight-reading and keyboard performance. *Science*, *257*(5066), 106–109. https://doi.org/10.1126/science.1621084

Shaffer, L. H. (1982). Rhythm and timing in skill. *Psychological Review*, *89*(2), 109–122. https://doi.org/10.1037/0033-295X.89.2.109

Shaffer, L. H. (1984). Timing in solo and duet piano performances. *The Quarterly Journal of Experimental Psychology*, *36*(4), 577–595. https://doi.org/10.1080/14640748408402180

Shaffer, L. H., Clarke, E. F., & Todd, N. P. (1985). Metre and rhythm in piano playing. *Cognition*, *20*(1), 61–77. https://doi.org/10.1016/0010-0277(85)90005-8

Shannon, C. E. and Weaver, W. (1949). *The Mathematical Theory of Communication,* Urbana, University of Illinois Press.

Sheridan, H., Maturi, K. S., & Kleinsmith, A. L. (2020). Eye movements during music reading: Toward a unified understanding of visual expertise. In K. D. Federmeier & E. R. Schotter (Eds.), *The psychology of learning and motivation: Gazing toward the future: Advances in eye movement theory and applications* (pp. 119–156). Elsevier Academic Press. https://doi.org/10.1016/bs.plm.2020.07.002

Silva, S., & Castro, S. L. (2019). The time will come: Evidence for an eye-audiation span in silent music reading. *Psychology of Music*, *47*(4), 504–520. https://doi.org/10.1177/0305735618765302

Simoens, V. L., & Tervaniemi, M. (2013). Auditory short-term memory activation during score reading. *PLOS ONE*, *8*(1), e53691. https://doi.org/10.1371/journal.pone.0053691

Simon, H., & Chase, W. (1988). Skill in chess. In D. Levy (Ed.), *Computer chess compendium* (pp. 175–188). Springer. https://doi.org/10.1007/978-1-4757-1968-0_18

Sloboda, J. (1974). The eye-hand span-an approach to the study of sight reading. *Psychology of Music*, *2*(2), 4–10. https://doi.org/10.1177/030573567422001

Sloboda, J. (1978). The psychology of music reading. *Psychology of Music*, *6*(2), 3–20. https://doi.org/10.1177/030573567862001

Sloboda, J. (2005). *Exploring the musical mind: Cognition, emotion, ability, function*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198530121.001.0001

Sloboda, J. A. (1976a). Visual perception of musical notation: Registering pitch symbols in memory. *The Quarterly Journal of Experimental Psychology*, *28*(1), 1–16. https://doi.org/10.1080/14640747608400532

Sloboda, J. A. (1976b). The effect of item position on the likelihood of identification by inference in prose reading and music reading. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *30*(4), 228–237. https://doi.org/10.1037/h0082064

Sloboda, J. A. (1977). Phrase units as determinants of visual processing in music reading. *British Journal of Psychology*, *68*(1), 117–124. https://doi.org/10.1111/j.2044-8295.1977.tb01566.x

Sloboda, J. A. (1984). Experimental studies of music reading: A review. *Music Perception*, *2*(2), 222–236. https://doi.org/10.2307/40285292

Sloboda, J. A., Clarke, E. F., Parncutt, R., & Raekallio, M. (1998). Determinants of finger choice in piano sight-reading. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(1), 185–203. https://doi.org/10.1037/0096-1523.24.1.185

Snyder, B. (2000). *Music and memory: An introduction*. MIT press.

Solomon, M. (1995) *Mozart: A life*. HarperCollins

Soulez, F., Rodet, X. & Schwarz, D. (2003). Improving polyphonic and poly-instrumental music to score alignment. In *4th International Conference on Music Information Retrieval*, Baltimore, U.S.A.

Starr, M. S., & Rayner, K. (2001). Eye movements during reading: Some current controversies. *Trends in Cognitive Sciences*, *5*(4), 156–163. https://doi.org/10.1016/S1364-6613(00)01619-3

Temperley, D. (2007). *Music and probability*. Mit Press.

Thompson W. B. (1985). *Sources of individual differences in music sight-reading skill* (Unpublished doctoral dissertation). University of Missouri, Columbia.

Thompson, S., & Lehmann, A. C. (2004). Strategies for sight-reading and improvising music. In A. Williamon (Ed.), *Musical excellence: Strategies and techniques to enhance performance* (pp. 143–160). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198525356.003.0008

Thompson, W. B. (1987). Music sight-reading skill in flute players. *The Journal of General Psychology*, *114*(4), 345–352. https://doi.org/10.1080/00221309.1987.9711072

Tinker, M. A. (1958). Recent studies of eye movements in reading. *Psychological Bulletin*, *55*(4), 215–231. https://doi.org/10.1037/h0041228

Todd, N. P. (1992). The dynamics of dynamics: A model of musical expression. *Journal of the Acoustical Society of America*, *91*(6), 3540–3550. https://doi.org/10.1121/1.402843

Townsend, B. G. (1991). *Relationships between sight-reading ability of college freshmen wind instrumentalists and music experience, band experience, and music aptitude* (Order No. 9214290). [Doctoral dissertation, The Pennsylvania State University]. ProQuest Dissertations and Theses Global.

Truitt, F. E., Clifton, C., Pollatsek, A., & Rayner, K. (1997). The perceptual span and the eye-hand span in sight reading music. *Visual Cognition*, *4*(2), 143–161. https://doi.org/10.1080/713756756

Tsangari, V. (2010). *An interactive software program to develop pianists' sight-reading ability* [Doctoral dissertation, University of Iowa]. https://doi.org/10.17077/etd.nqub1mfn

Veldre, A., & Andrews, S. (2015). Parafoveal preview benefit is modulated by the precision of skilled readers' lexical representations. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(1), 219–232. https://doi.org/10.1037/xhp0000017

Vidwans, A., Gururani, S., Wu, C. W., Subramanian, V., Swaminathan, R. V., & Lerch, A. (2017). Objective descriptors for the assessment of student music performances. In *Proceedings of 2017 Audio Engineering Society*

International Conference on Semantic Audio, Erlangen, Germany. https://www.aes.org/e-lib/browse.cfm?elib=18758

Waters, A. J., & Underwood, G. (1998). Eye movements in a simple music reading task: A study of expert and novice musicians. *Psychology of Music*, *26*(1), 46–60. https://doi.org/10.1177/0305735698261005

Waters, A. J., Townsend, E., & Underwood, G. (1998). Expertise in musical sight reading: A study of pianists. *British Journal of Psychology*, *89*(1), 123–149. https://doi.org/10.1111/j.2044-8295.1998.tb02676.x

Watkins, A., & Hughes, M. A. (1986). The effect of an accompanying situation on the improvement of students' sight reading skills. *Psychology of Music*, *14*(2), 97–110. https://doi.org/10.1177/0305735686142002

Weaver, H. E. (1943). A survey of visual processes in reading differently constructed musical selections. *Psychological Monographs*, *55*(1), 1–30.

White, S. J. (2008). Eye movement control during reading: Effects of word frequency and orthographic familiarity. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(1), 205–223. https://doi.org/10.1037/0096-1523.34.1.205

Williamon, A., & Valentine, E. (2002). The role of retrieval structures in memorizing music. *Cognitive psychology*, *44*(1), 1–32. https://doi.org/10.1006/cogp.2001.0759

Wolf, T. (1976). A cognitive model of musical sight-reading. *Journal of Psycholinguistic Research*, *5*(2), 143–171. https://doi.org/10.1007/BF01067255

Wolfs, Z. G., Boshuizen, H. P. A. (Els), & van Strien, J. L. H. (2020). The role of positional knowledge and tonal approaches in cellists' sight-reading.

Musicae Scientiae, *24*(1), 3–20. https://doi.org/10.1177/1029864918762269

Wöllner, C., Halfpenny, E., Ho, S., & Kurosawa, K. (2003). The effects of distracted inner hearing on sight-reading. *Psychology of Music*, *31*(4), 377–389. https://doi.org/10.1177/03057356030314003

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

Wong, Y. K., Peng, C., Fratus, K. N., Woodman, G. F., & Gauthier, I. (2014). Perceptual expertise and top–down expectation of musical notation engages the primary visual cortex. *Journal of Cognitive Neuroscience*, *26*(8), 1629–1643. https://doi.org/10.1162/jocn_a_00616

Wristen, B. (2005). Cognition and motor execution in piano sight-reading: A review of literature. *Update: Applications of Research in Music Education*, *24*(1), 44–56. https://doi.org/10.1177/87551233050240010106

Wristen, B., Evans, S., & Stergiou, N. (2006). Sight-reading versus repertoire performance on the piano: A case study using high-speed motion analysis. *Medical Problems of Performing Artists*, *21*(1), 10–16. https://doi.org/10.21091/mppa.2006.1003

Wurtz, P., Mueri, R. M., & Wiesendanger, M. (2009). Sight-reading of violinists: Eye movements anticipate the musical flow. *Experimental Brain Research*, *194*(3), 445–450. https://doi.org/10.1007/s00221-009-1719-3

Young, L. J. (1971). *A study of the eye-movements and eye-hand temporal relationships of successful and unsuccessful piano sight-readers while* *piano sight-reading* (Unpublished doctoral dissertation). Indiana University.

Youngblood, J. E. (1958). Style as information. *Journal of Music Theory*, *2*(1), 24–35. https://doi.org/10.2307/842928

Zhukov, K. (2014). Evaluating new approaches to teaching of sight-reading skills to advanced pianists. *Music Education Research*, *16*(1), 70–87. https://doi.org/10.1080/14613808.2013.819845

Zhukov, K. (2017). Experiential (informal/non-formal) practice does not improve sight-reading skills. *Musicae Scientiae*, *21*(4), 418–429. https://doi.org/10.1177/1029864916684193

Zhukov, K., Khuu, S., & McPherson, G. E. (2019). Eye-movement efficiency and sight-reading expertise in woodwind players. *Journal of Eye Movement Research*, *12*(2), 10.16910/jemr.12.2.6. https://doi.org/10.16910/jemr.12.2.6

Zhukov, K., & McPherson, G. (2022). Sight-reading. In G. E. Mcpherson (Ed.), *The Oxford handbook of music performance: Volume 1* (pp. 192–213). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190056285.013.11

Zhukov, K., Viney, L., Riddle, G., Teniswood-Harvey, A., & Fujimura, K. (2016). Improving sight-reading skills in advanced pianists: A hybrid approach. *Psychology of Music*, *44*(2), 155–167. https://doi.org/10.1177/0305735614550229

국문 초록

본 논문은 전문 연주자들 사이에서 나타나는 초견 능력의 개인차를 눈-손 간격(EHS)을 중심으로 탐구한다. 초견이란 사전 연습 없이 악보를 처음 봄과 동시에 바로 연주하는 것을 의미하며, 연주자들의 탁월한 음악적 능력을 대표하는 능력이자 음악가들이 반드시 습득해야 하는 가장 기본적인 능력으로 알려져 있다. 흥미로운 점은 악기를 다루는 기술이 뛰어난 전문 연주자들 사이에서도 초견 능력은 제각각이라는 점이다. 본 논문은 이러한 현상에 주목하여 다음의 연구 질문을 다룬다: 왜 특정 연주자들은 초견에 능숙하고 다른 연주자들은 어려움을 겪는가?

본 논문은 전문 피아니스트들을 대상으로 그들의 눈-손 간격을 측정하고 그것과 연주 정확도 측면에서 평가된 초견 능숙도와의 상관관계를 조사하였다. 눈-손 간격은 초견 연주 시 악보에서 눈이 읽고 있는 지점과 손이 연주하고 있는 지점 사이의 간격으로, 역사적으로 음악심리학 분야에서 초견 능숙도의 절대적인 지표로 알려져 왔다. 그러나 눈-손 간격과 초견 능숙도 간의 관계에 음악적 변수를 고려한 통합적인 관점은 상대적으로 부족하였다. 본 논문은 이를 고려하여 초견에 대한 세 가지 영역을 설정하고 그들 사이의 상호 관계를 탐구하였다. 각 영역의 지표에는 음악의 복잡성과 연주 템포 (음악 영역), 눈-손 간격 (인지 영역), 그리고 연주 정확도 (행동 영역)이 포함되었다.

실험에는 서른한 명의 전문 피아니스트들이 참여하였고, 그들은 각각 두 가지 다른 복잡성과 연주 템포로 구성된 네 개의 악곡을 초견으로 연주하였다. 본 논문은 참가자들의 눈-손 간격을 측정하고, 그들의 연주 정확도를 평가하였으며, 연주 정확도 값에 따라 참가자들을 세 그룹으로 분류하고 각 그룹에 따른 눈-손 간격의 변화와 음악 영역의 영향을 조사하였다. 연구 결과, 놀랍게도 참가자들의 눈-손 간격은 연주 정확도에만 근거하여 변하지 않았다. 대조적으로, 눈-손 간격과 연주 정확도의 관계는 음악의 복잡성과 연주 템포의 조합인 초견 과제의 난이도에 따라 계속해서 변하는 것으로 나타났다. 특히 연주 정확도가 높은 연주자일수록 이러한 가변성을 더욱 활용하였다. 이는 초견에 능숙한 연주자들은 그렇지 않은 연주자들보다 항상 긴 눈-손 간격을 가지는 것이 아니라 음악적 특성과 같은 중개 요인에 따라 그들의 눈-손 간격을 유연하게 조절해 가며 초견에 임한다는 것을 보여준다. 종합하자면, 본 논문은 초견 능숙도의 절대적인 지표로 여겨졌던 눈-손 간격이 사실은 인지적 전략이라는 인식의 전환을 보여준다. 본 논문은 눈-손 간격이 초견 능숙도에 단순 비례하는 것이 아니라 음악적 맥락에 따라 조정되는 유연한 전략임을 밝혀낸 점에서 의미가 있다.

초견에 대한 과학적 탐구는 다양한 측면에서 시사점을 제공한다. 인지 과학의 관점에서 초견에 관한 실증적 연구는 인간의 다중 감각의 통합 및 처리 그리고 주의, 기억, 행동, 예측과 같은 고차원적인 인지 기능의 조정에 대한 이해를 심화시킨다. 능숙한 초견에 대한 실증적

연구는 또한 인간의 전문성과 숙련된 행동이 어떻게 습득되고, 개발되고, 내재화되는지에 대한 이해의 창을 제공한다. 교육학적 관점에서, 초견의 인지적 메커니즘을 조사하는 것은 교육자들이 근거에 기반한 교수법과 교육 환경을 마련하고 학습자들의 개별적인 도전과제를 고려한 맞춤형 초견 훈련 프로그램을 고안하는 데 도움이 될 수 있다. 음악적 관점에서 초견에 관한 체계적인 탐구는 역사적으로 신비롭게 여겨지던 뛰어난 초견 능력이 구체적으로 어떠한 종류의 능력인지를 밝히며 여전히 많은 전문 연주자에게 고민거리인 초견 능력을 전략적으로 연마하는 것에 기여할 것이다.

키워드: 음악 인지, 음악 연주, 초견, 피아니스트, 시선 추적, 눈-손 간격, 복잡도, 연주 템포
학번: 2018-34350

Acknowledgments

부족함이 많은 제가 학위 논문을 완성하고 박사학위를 취득할 수 있게 된 것은 많은 분들의 도움 덕분입니다. 이 자리를 빌려 감사한 분들께 짧게나마 감사의 인사를 올리고자 합니다.

먼저 학사부터 박사까지의 여정 동안 훌륭한 스승이자 인생의 멘토로 저를 이끌어 주신 이석원 교수님께 진심으로 감사드립니다. 이석원 교수님께 배울 수 있었던 것은 제 인생 최고의 행운이자 행복이었습니다. 어리석은 제자는 그동안 걱정 없이 연구할 수 있던 것이 자신이 잘해서가 아니라 모두 당신께서 풍족한 환경을 제공해 주시고 부족한 제자를 넓은 아량과 인내로 지도해주셨기 때문이었음을 이제야 깨닫습니다. 교수님의 가르침을 잊지 않고 계속해서 겸손한 자세로 학문에 정진하겠습니다.

박사학위 논문의 심사를 맡아주시고 논문을 개선하는 데 큰 도움을 주신 심사위원 교수님들께도 진심으로 감사드립니다. 이른 전공에 입학한 열여덟 새내기 시절부터 저를 알뜰살뜰하게 챙겨주시고 학문적으로도 귀감이 되어주신 오희숙 교수님, 미숙한 학부생이 음악심리학 분야에서 꿈을 갖고 열심히 연구할 수 있도록 격려해 주시고 영감의 원천이 되어주신 이경면 교수님, 여러 연구의 공동지도를 맡아주시며 연구자로 성장하는 길목에서 빛을 비추어주신 이교구 교수님, 연구실의 든든한 선배로 동고동락하며 학문적으로나 개인적으로나 항상 따뜻하게 조언해 주시고 동기부여가 되어주신 박정미 교수님, 그리고 음악학계의 원로로서 존경해 마지않는 허영한 교수님께 진심으로 감사드립니다. 제 박사학위 논문의 심사위원으로 교수님들을 모실 수 있어 대단히 영광이었습니다.

대학원 기간 수행한 초견 연구에 대해 귀중한 조언을 해주시고 박사학위 논문의 기반이 되었던 학술지 논문이 출판되기까지 다방면으로 도움을 주신 정천기 교수님과 김연 교수님께도 깊이

감사드립니다. 또한 다년간 저의 연구에 대해 함께 고민하며 연구를 발전시키는 데 실질적인 도움을 주신 유승연 박사님께도 진심으로 감사드립니다.

학위 논문을 쓰는 여정에서 개인적으로 힘이 되어주신 분들께도 감사 인사를 전하고 싶습니다. 부산에서 배움의 기회가 필요할 때 따뜻하게 받아주시고 환대해 주신 주성준 교수님과 시각 및 읽기 연구실 선생님들께 진심으로 감사드립니다. 육아와 학업이라는 공통의 길을 먼저 걸어간 선배로서 제가 힘이 들 때마다 용기를 북돋아 준 김나연 박사님과 저에게 현명한 조언을 아끼지 않으며 일상의 지원군이 되어준 최인영 선생님께도 깊이 감사드립니다. 늘 한국 딸이 최고라며 무한한 응원과 사랑을 보내주시는 미국의 양부모님 Joel Popkin 과 Zenie Popkin 께도 감사드립니다. 사랑하는 나의 친구들에게도 고마움을 전합니다. 이곳에는 미처 담지 못하였지만, 지금까지 저를 응원해 주시고 격려해 주신 모든 분께 진심으로 감사드립니다.

마지막으로 사랑하는 가족에게 감사의 마음을 전합니다. 피아노로 시작하여 음악심리학으로 끝을 맺는 여정 동안 딸이 어떠한 선택을 하든 늘 응원해 주시고 물심양면으로 지원해 주신 부모님께 진심으로 감사드립니다. 부모님의 사랑과 헌신이 있었기에 지금까지의 여정이 가능했습니다. 제가 어머니의 도움을 받아 아기를 키우고 있듯이 또한 그 시절 당신의 딸을 도우며 저를 사랑으로 키워주신 외할머니와 하늘에 계신 외할아버지께도 감사합니다. 며느리가 걷는 학자의 길을 지지해 주시고 늘 무한한 사랑과 힘을 주시는 시부모님께도 진심으로 감사드립니다. 항상 밝은 미소로 응원해 주시는 문옥 아가씨, 재준씨, 그리고 지수에게도 깊이 감사합니다.

바쁜 엄마와 살면서도 그저 건강하고 밝게 자라준 사랑하는 나의 아들 황민상에게 진심으로 고맙습니다. 그리고 지금까지의 여정에서 나에게 끊임없이 용기를 실어주며 든든한 버팀목이자 안식처가 되어준 나의 남편 황문원, 가슴 깊이 사랑하고, 존경하고, 감사합니다.