음악학석사 학위논문

A Cognitive Model of Sight-Reading as the Eye-Ear-Hand Collaborative Processes

눈-귀-손 협력 과정으로서의 초견 연주 인지 모델

2018년 2월

서울대학교 대학원

음악과 이론전공

임 여 은

A Cognitive Model of Sight-Reading as the Eye-Ear-Hand Collaborative Processes

지도교수 이 석 원

이 논문을 음악학석사 학위논문으로 제출함
2018년 2월

서울대학교 대학원
음악과 이론전공
임 여 은

임여은의 석사 학위논문을 인준함
2018년 1월

위 원 장 ______ 민 은 기 ______ (인)

부위원장 ______ 오 희 숙 ______ (인)

위 원 ______ 이 석 원 ______ (인)

Abstract

Yeoeun Lim
Department of Music
The Graduate School
Seoul National University

Musical performance is a real-world activity that is amenable to behavioral and cognitive fractionation and investigations into its physiological basis. This thesis concerns one basic but important musical ability – sight-reading. Because sight-reading involves complex visual, auditory, and motor processing tasks that occur simultaneously, an investigation of the cognitive mechanisms underlying it, is likely to contribute a fundamental understanding of these processes both within and outside the musical field.

This thesis aims to examine the neurological and physiological mechanisms of sight-reading from theoretical perspectives and explore factors that can influence these mechanisms, based on empirical evidence. Understanding the neurological mechanisms of sight-reading can identify particular brain areas that integrate the visual, motor, and auditory domains. Exploring the physiological mechanisms of sight-reading focuses on eye movements and can provide empirical evidence for the neurological specificity of sight-reading. This thesis additionally deals with the various factors that engage sight-reading mechanisms and examines the influence of several variables on sight-reading proficiency, considering both subjective and objective factors.

Finally, the present thesis presents a series of experiments to point out and characterize technical and musical limitations, which were lacking in previous studies, and test a modified paradigm. In particular, this thesis proposes a new measurement of musical complexity and objectively investigates its influence on sight-reading proficiency through a behavioral experiment. Furthermore, this thesis examines a correlation between performance accuracy and the eye-hand span, which represents the physiological mechanisms of sight-reading, and investigates the influence of variable factors on the mechanisms through a physiological experiment.

Keywords : Sight-reading, Eye tracking, Eye-hand Span, Complexity, Entropy, Dynamic Time Warping
Student Numbers : 2016-21879

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

1.1 Research Background

Among various musical aptitudes, sight-reading is one of the most important performance skills (Kopiez and Lee, 2008; 2006; Kopiez et al., 2006). McPherson (1995; 1993), for example, classified sight-reading as one of the five distinct performance skills: performing rehearsed music, sight-reading, playing from memory, playing by ear and improvisation, suggesting the use of independent strategies from others during sight-reading. Sight-reading requires not only excellent instrumental techniques, but also an extensive knowledge of stylistic musical features because performers rapidly convert symbolic music notation into motor execution under strict temporal regulation (Kopiez et al., 2006). In this regard, the level of musicianship, such as musical expressivity or instrumental technique, are distinct from sight-reading skills, even among professional pianists (Wolf, 1976).

Indeed, sight-reading is complex task that simultaneously involves visual, auditory, and motor capabilities (Schön et al., 2002; Sergent et al., 1992). Because the eyes of a performer are on a score, the hands must move continuously, the ears must both hear sounds, and the mind must memorize what is seen during sight-reading. Theses factors have attracted steady attention from musical psychologists interested in sight-reading skills (Lehmann and Kopiez, 2009). Additionally, numerous studies have attempted to explore the mechanisms of sight-reading by applying it to scientific analyses. Chronologically, early sight-reading experiments started in the 1940s and multidimensional approaches to sight-reading performance began after the standardization of sight-reading evaluations was introduced in the

1970s. Educational discussions based on the theoretical and empirical implications increased throughout the 1990s, and over the last decade neurological studies that include brain imaging techniques have been conducted to investigate the neurological mechanisms of sight-reading.

1.2 Purpose of This Study

This study aims to examine the neurological and physiological mechanisms underlying sight-reading from theoretical perspectives and explore variable factors that influence these mechanisms based on empirical evidence. In particular, this study will point out and characterize technical and musical limitations, which were absent from past studies, and ultimately suggests a new cognitive model of sight-reading. Additionally, sight-reading represents a real-world activity that is amenable to behavioral and cognitive fractionation, so investigations into its physiological basis and the cognitive mechanisms underlying it are likely to contribute a fundamental understanding of these processes both within and outside the musical field.

1.3 Study Outline

The remainder of this thesis is organized as follows. Chapter 2 reviews the existing neurological and physiological studies on the mechanisms of sight-reading. Specific brain areas associated with sight-reading performance are discussed. Particularly, neuroimaging studies that compare brain activities between sight-reading and nonmusical tasks provide neurological evidence for the mechanisms underlying sight-reading. Eye movements are presented as the index of the physiological mechanisms of sight-reading. Additionally, the factors that influence sight-reading performance are discussed in Chapter 2.

Chapters 3 and 4 present a series of experiments that were carried out to complement the limitations of the previous studies and test a modified paradigm. Finally, Chapter 5 summarizes the main findings of Chapters 3 and 4 and places the results in the context of previous studies of sight-reading. A cognitive model of sight-reading is proposed that is based on past and present empirical work. The limitations of the experiments and plans for further experiments are also described in Chapter 5.

Chapter 2. Literature Review

This chapter discusses the mechanisms involved in sight-reading from a neurological and physiological perspective on a theoretical level. Understanding the neurological mechanisms of sight-reading thereby can identify particular brain areas that integrate the visual, motor, and auditory domains. The physiological mechanisms of sight-reading focus on eye movements and provide empirical evidence for the neurological specificity of sight-reading. Finally, this chapter deals with the various factors that engage sight-reading mechanisms and examines the influence of several variables on sight-reading proficiency, considering both subjective and objective factors.

2.1 Mechanisms Underlying Sight-Reading

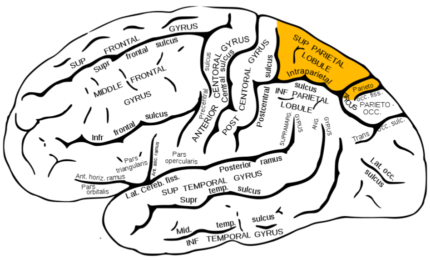
2.1.1. Neurological Perspectives

Two neuroimgaging studies that investigated areas of brain activation during sight-reading showed that the existence of neural subsystems specialized in music reading and that such neural networks are adjacent, but distinct from those involved in text reading. In the first neuroimaging study of sight-reading (Sergent et al., 1992), the authors asked ten professional pianists to sight-read an unfamiliar partita written by J. S. Bach. They measured areas of brain activation using positron emission tomography (PET).[1] In addition to the experimental conditions for sight-reading, six

1) PET is a method that involves injecting a radiolabeled biological molecule into the bloodstream to obtain information about metabolic processes in a particular part

other conditions were used as a control to separate the main task from the component operations. The control conditions included simple visual fixation of the screen, listening to ascending and descending scales, playing ascending and descending scales on the keyboard with auditory feedback, viewing a single dot on the screen, reading a musical score without performance and reading a musical score with listening to its performance. The experimental results indicated that the superior parietal cortex (SPC), which is generally not considered to be involved in verbal operations, was only activated when simultaneously reading, playing, and listening to the musical score. The SPC is located slightly behind the crown and acts to convert spatial information into bodily movements. For example, it manages the extraction of spatial features related to musical notation (Stewart, 2005) and mediates sensorimotor transformations for trained actions with visual feedback and finger positioning (Schön et al., 2002).

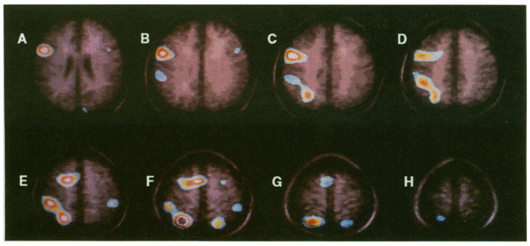
Figure 1. Superior Parietal Lobule (highlighted in orange)



---

of the body and to identify tissues that have absorbed the molecules with a PET scan of radiation release. PET facilitates investigations of functional characteristics, rather than the structure of the brain. While the spatial and temporal resolution of PET are lower than those of functional magnetic resonance imaging (fMRI), PET can be used to visualize the metabolism of specific neurotransmitters.

Figure 2. Cortical activation during sight-reading and piano performance.

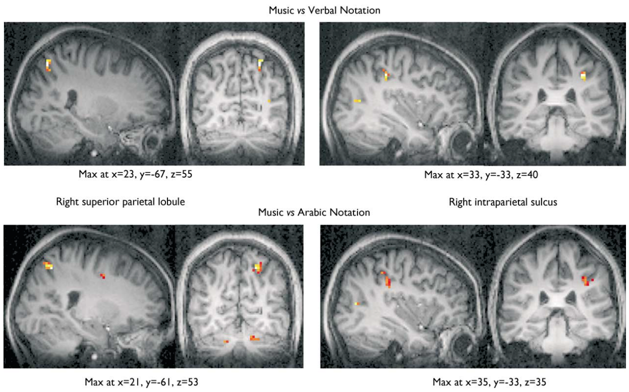


Another neuroimaging study measured the activation of brain areas for music, verbal and number notations using fMRI[2] (Schön et al., 2002). The authors also demonstrated distinctive areas of SPC activation while reading musical notes. In this study, nine pianists were assigned to three different experimental conditions: playing the five musical notes, and making verbal and Arabic notations on a keyboard. For music notation, two major foci of activation were observed in the right SPC and intraparietal sulcus (IPS). Specifically, these two areas were significantly more active while reading music notation compared to verbal or numerical notations. Similar to the SPC, the IPS is the sulcus between the superior and inferior parietal lobules. It mainly involves sensorimotor integration (Jäncke et al., 2001) and visuospatial processing (Simon et al., 2002). These results suggest the involvement of well-defined right parietal regions such as the SPC or IPS, in processing musical notation.

---

2) fMRI measures brain activity by detecting changes associated with blood flow. This technique relies on the principle that cerebral blood flow and neuronal activation are coupled. When a brain area is in use, blood flow to that region also increases. Particularly, fMRI uses blood-oxygen-level dependent (BOLD) contrast to observe regional differences in cerebral blood flow to delineate regional activity. In music research, BOLD imaging is generally used to determine areas of the brain that are responsible for music reading.

Figure 3. Comparison of neural activation between music and verbal notation (above) and music and number notation (below). Images were acquired by subtracting verbal and number notations from music notation. The red and yellow colors represent higher signals for the music reading task relative to the verbal or number notation reading tasks.



Taken together, music notation becomes coded as spatial position during sight-reading, whereas text or numbers are coded according to their form. The conversion of visual information into spatial motion underlies the neurological basis for sight-reading.

2.1.2. Physiological Perspectives

What factor accounts for the neurological characteristics of sight-reading that are distinct from other cognitive activities? It is the eye. A musical score markedly limits eye movements. The eye moves more precisely when reading music than when viewing visual information, such as moving images. Such a system that restricts eye movements uses a mechanism that is discrete from conventional visual processing. Eye movements can be explained by fixations and saccades while viewing visual information. Fixations are the pausing of eye movements (200-300 milliseconds), during which information processing and cognitive activities occur and the information reaches the brain (Madell and Héébert, 2009). Saccades are extremely rapid eye movements that occur at intervals of approximately 40-50 milliseconds. Incoming information is restricted and rarely processed during saccades. Most fixations usually, but not always have directionality going forward. For example, when reading long, complex musical scores, the eye retrogresses to read the score again to clarify the phrase (Goolsby, 1994a; 1994b). In addition, a longer gaze fixation time is associated with more complicated information processing and increased difficulty in recognizing objects (Goolsby, 1994b; Rayner, 1998; Pollatsek et al., 1986). Therefore, observing the fixations and saccades of performers while reading music allows for an understanding of how notes are processed differently in the brain in comparison with other types of symbol systems, such as letters and numbers.

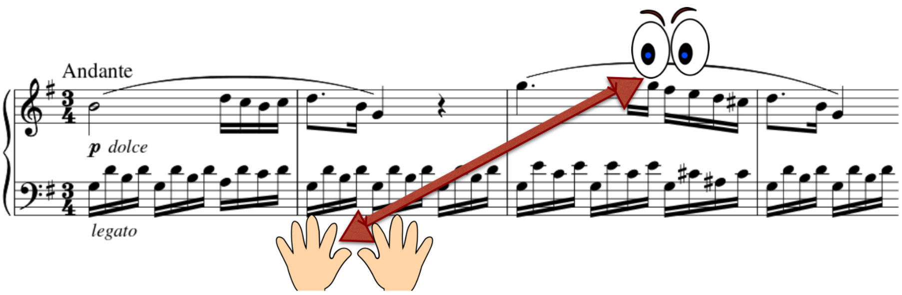
Performers' eyes are constantly moving during sight-reading. By measuring their eye movements using an eye tracker, how the eye moves on a musical score can be investigated. For example, Figure 4 shows a simplified representation of the eye movements of a skilled sight-reader.

Figure 4. Example of the fixations and saccades of a skilled sight-reader during sight-reading (Penttinen, 2013). Numbers above notes indicate fixations in which the eye stopped, and arrows represent saccades.

Figure 4 suggests that the eye does not look at every single note, but instead reads several notes at the same time. In other words, professional pianists perceive, process, and memorize multiple notes simultaneously whenever their gaze is stopped. Indeed, from the third to the fourth fixation in Figure 4, the player skipped two or three notes. Thus, professional performers do not only look at a single note in a single fixation, but also see an average of five to twelve notes (or sometimes even more) together according to the musical style during sight-reading. The number of notes that are processed while the gaze stops is termed the perceptual span (Gilman and Underwood, 2003; Truitt et al., 1997; Rayner, 1993). The better the sight-reader, the larger the span, so the best players can read more notes in one fixation.

On the other hand, a series of coordination processes from the eye movements to motor execution is exceedingly important because sight-reading is an operation in which perception and physical movement both occur within a tremendously short interval. The eye-hand span (EHS) represents the link from the perceptual span to motor responses. The EHS is the distance between the hand and eye position on the score and is related to working memory or working memory capacity (Furneaux and Land, 1999; Kinsler and Carpenter, 1995). It can be measured in three ways: in note or beat (the number of notes or beats between the hand and eye), or in time (the length of time between fixation and performance).

Figure 5. Visualization of the EHS



Earlier studies of the EHS before the development of eye-tracking techniques were conducted by either switching off the light or the music slide. In this way, Sloboda (1974) found that the average span of all sight-readers was 3.8 notes, and the span of the best sight-readers was 6.8 notes. As the technology for eye-tracking recording has improved, many studies have investigated the lengths of the EHS (Table 1).

Table 1. Summary of the EHS literature.[3)]

Study		Eye tracker	EHS			Participants		Tempo	Stimuli
			note	time	beat	type	n		
1	Weaver (1943)	O	1.9-3.1			musicians	15	X	single melody, dual-staved music
2	Sloboda (1974)	X	3.8-6.8			musicians (but various)	10	O	single melody
3	Goolsby (1994b)	O			4	skilled/less skilled musicians	2	△	single melody
4	Truitt et al. (1997)	O			1-2	pianists	8	△	single melody
5	Furneaux and Land (1999)	O	2-4	1s		novice, intermediate, professional	8 (3,3,2)	X	dual-staved music
6	Gildman and Underwood (2003)	O			3/4-1	pianists	40	X	Bach Chorales
7	Wurtz et al. (2009)	O	3-6	1s		violinists	7	X	single melody (Corelli, Telemann)
8	Rosemann et al. (2015)	O		1-1.5s	0.5±	pianists	9	O	Bach Flute Sonata
9	Cara (2017)	O	2.71-7.34		1.36-4.71	pianists	22	X	Ligeti Piano Etude
Total			1.9-7.34	1-1.5s	0.5-4.71				

In most cases, the EHS has been found to vary according to sight-reading proficiency. For example, skilled pianists tended to have a large EHS (Cara, 2017; Gilman and Underwood, 2003; Furneaux and Land,

3) **Eye tracker**: Whether to use an eye tracker; **EHS**: The value of the EHS in note, beat, and time indexes; **Participants**: The number and type of subjects who took part in experiments; **Stimuli**: The type of sight-reading materials used in the experiments; **Tempo**: X and O indicate the inconsistency and consistency, respectively, of performance tempo among the participants, △ indicates that the metronome was given to participants before the sight-reading task but was removed during sight-reading.

1999; Truitt et al., 1997; Sloboda, 1974; Weaver, 1943). These observations raise the question of how the better sight-readers achieve a longer EHS?

Human has a limited memory capacity. By reducing the size of information through chunking, we can process and memorize more information than our capacity. Chunking is a grouping of listed information into meaningful units. In Letter 1, it would be hard to remember if one tried to memorize the letters in alphabetical units.

Letter 1

'sightreadingandtheeyehandspan'

Instead, if you put these twenty-nine pieces of information into four groups, such as "sight-reading" and "the eye-hand span" so that memorization would be simplified (Letter 2).

Letter 2

'sight-reading and the eye-hand span'

It is the same with music. Less-skilled sight-readers have a shorter EHS because their eyes follow all the notes, and they play with a large amount of information during sight-reading. In contrast, in individuals with better sight-reading, information reduction will be achieved by chunking, and the notes of the score will be recognized as patterns. Skilled sight-readers, therefore, play with relatively little information, and the most proficient sight-readers can recognize the bigger musical patterns on a score.

In sum, the EHS indicates how long visual information is stored in a buffer before the output into finger movements, and represents an objective and efficient indicator of the difference between skilled and less-skilled sight-readers.

2.2 Variable Factors

The previous section examined the neurological and physiological mechanisms underlying sight-reading performance. In this section, the major variable factors that influence the mechanisms underlying sight-reading are discussed. As sight-reading proficiency varies depending upon both the input stimulus of the score and individual differences, the factors can be divided into two categories: subjective and objective variables.

2.2.1. Subjective Variables

Many EHS studies have shown that the length of the EHS varies according to the skill level of sight-reading (Cara, 2017; Gilman and Underwood, 2003; Furneaux and Land, 1999; Truitt et al., 1997; Sloboda, 1974; Weaver, 1943). Specifically, skilled sight-readers had a larger EHS than less skilled sight-readers. As mentioned above, Sloboda (1974) demonstrated that good sight-readers had a roughly three-note larger EHS than average sight-readers. Truitt et al. (1997) also showed that the span of the skilled sight-readers was approximately one beat larger than that of less skilled sight-readers. Similarly, the EHS of poor sight-readers was significantly smaller and extended to less than a beat in the study of Gilman and Underwood (2003). Cara (2017) also found that expert pianists had a larger EHS, and the difference in the value of the EHS that depended upon the skill level was the greatest in his study compared to other published reports. The range of the EHS covered approximately three to eight notes and two to five beats.

Why does the length of the EHS differ depending upon the subject of sight-reading? The factors that influencing sight-reading proficiency can be divided into two main categories: practice-dependent and -independent factors. Practice-dependent factors are associated with musical experience, including years of practice. Notably, the accumulated amount of time spent on activities related to sight-reading is important for sight-reading achievement. In contrast, practice-independent factors are related to general and elementary cognitive skills. Intriguingly, the size and degree of the effects of these factors differ. Kopiez and Lee (2006) investigated the predictors of successful sight-reading performance and ranked these predictors. The best predictors were trilling speed, sight-reading expertise up to the age of 15, speed of information processing and inner hearing. Hence, sight-reading proficiency could be determined by a combination of practice-dependent factors, such as sight-reading expertise and inner hearing, as well as practice-independent predictors, such as trilling speed and the speed of information processing.

2.2.2. Objective Variables

Several musical parameters, including texture (Weaver, 1943), complexity (Cara, 2017; Rosemann, 2015; Wurtz, 2009; Kopiez and Lee, 2006) and tonality (Podolak, 2013), have been examined as objective variables that may influence the mechanisms underlying sight-reading. Among the objective variables, the significance of the musical complexity was emphasized the most because pattern recognition readily changes according to the complexity of music. Kopiez and Lee (2008; 2006) suggested that the combination of the predictor variables for sight-reading proficiency varied along with the complexity level of music. For example, when the sight-reading material had a low complexity, practice-dependent factors, such as general pianistic expertise, were important for the success of sight-reading performance. However, when the sight-reading material had the highest level of complexity, practice-independent factors, such as psychomotor speed and the speed of information processing, became the dominant predictors, even though practice-dependent factors still remained important.

Moreover, musical complexity can also affect the eye movements and the length of the EHS. A piece with higher complexity is associated with lower anticipation of notes, longer fixation duration and a tendency for more regressive fixations (Wurtz et al., 2009). Similarly, the number of beats was significantly higher for low complexity music than for high complexity music (Rosemann et al., 2015). Finally, Cara (2017) indicated that an intricate musical structure or increased complexity of music had a negative effect on the length of the EHS.

Chapter 3. Behavioral Experiment

3.1 Aims

Although it is obvious that musical complexity affects the physiological mechanisms underlying sight-reading, some points remain to be investigated. Indeed, previous studies used different criteria for complexity, so it is difficult to effectively summarize previous findings. Furthermore, only quantitative properties of music, such as the total number of notes or the total duration of the piece, were considered in the classification of the level of complexity. However, qualitative properties of music, such as pitch distribution or note duration difficulty, are also critical variables that affect musical complexity. For example, the level of complexity was divided according to the number of notes or the total duration of sight-reading material in the previous studies (Wurtz et al., 2009; Kopiez and Lee, 2008; 2006;). Yet, in many cases performers do not solely judge the level of complexity based on the quantitative factors. If there are two pieces with the same number of notes or total duration, the level of complexity will differ by qualitative characteristics of music, such as the chromatic distributions of pitch-classes or the rhythmic value that is used. Thus, musical complexity can according to both the quantity and quality of music.

The rationale for this behavioral experiment was based on the lack of previous consideration of qualitative differences in the classification of the level of complexity. Because it is challenging to define and quantify the quality of musical complexity, qualitative characteristics of music, such as pitch and rhythm difficulties, were excluded as criteria of complexity. Therefore, this study had three main aims. First, it aimed to propose a new measurement of musical complexity defined in terms of both qualitative and quantitative manners. In particular, I suggest pitch and note duration as criteria for complexity and examine the reliability of sight-reading materials using an existing tool, such as entropy. Second, this study aimed to evaluate sight-reading proficiency in an objective manner to allow for a more rigorous investigation of the influence of musical complexity on sight-reading accuracy. Dynamic time warping was conducted for objective evaluations of sight-reading. Finally, this study aimed to investigate the influence of musical complexity on sight-reading proficiency.

3.2 Methods

3.2.1 Participants

A total of 32 piano major students (28 females and 4 males; age range 18-24 years, mean age=20.17 years) who majored in piano at Seoul National University participated in this study. On average, participants had begun their piano training at 5.93 years (SD=0.923) of age, and had played the piano for 15.11 years (SD=2.149). They were thus regarded as having a high level of expertise in piano performance and were familiar with sight-reading.

3.2.2 Materials

Standards of Musical Complexity

To better reflect the qualitative aspects of music related the level of complexity and to complement the limitations of the previous studies, pitch and note duration were used as standards of complexity. Because pitch and rhythm have been regarded as fundamental elements in the history of Western classical music, this study used those two musical elements as determinants of musical complexity. The level of complexity was determined according to difficulties in pitch and duration, so the overall complexity of the music was directly proportional to the levels of pitch and durational difficulties. For example, the more accidentals that notes had, the higher was the level of complexity of the piece in terms of pitch difficulties. Similarly, the shorter the duration that the notes had, the higher the level of complexity of the piece in terms of note duration difficulty.

Composition of Sight-reading Materials

Sight-reading materials were composed in a quantitative manner, and the level of complexity was objectively differentiated by comparing the materials quantitatively. Three musical pieces with three different levels of complexity (i.e., low, medium and high) were composed specifically for this experiment. To prevent the intervention of unexpected variables other than musical complexity, each piece had the same key (C Major), time signature (4/4 meter), length (16 measures, 48 seconds), and tempo (80 beats per minute).

Among the three levels, the lowest complexity pieces had a total of 19 half notes and over half notes, thereby making the low level the simplest in terms of duration. Additionally, there were no accidentals, so the low complexity pieces had the simplest complexity level in terms of pitch. The total number of notes used in the low level complexity piece was 130 (Figure 6). The piece at the medium level of complexity had eighth half notes and over half notes, which were halved compared to the low level. Accidentals were now included. The number of accidentals was one to two notes per bar, so each bar had at least one accidental and the total number of accidentals was 27. A total of 160 notes were used at the medium level, which was 30 more notes than the low-level piece. This same difference of

30 notes was also applied to the high level of complexity (Figure 7).

Finally, half and over half notes, which had the longest duration, disappeared entirely in the high level of complexity piece. A new durational value, the sixteenth note, was introduced so that a total of twelve sixteenth notes were used in the high-level. The longest duration was 19 notes in the low, 8 notes in the medium, and 0 in the high level, respectively. The level of difficulty among the complexities was equally controlled. Accidentals appeared at 3 to 8 instances per bar, bringing the total number of accidentals to 84, a three-fold increase compared with the 27 accidentals in the medium level piece. Therefore, the level of complexity in regard to pitch was the highest. In an effort to maintain objectivity, the total number of notes increased from 160 to 190, a 30-note increase, which was comparable to the 30-note increase from the low to medium pieces (Figure 8). In summary, Table 2 shows the quantitative composition of the sight-reading materials.

Figure 6. Low complexity

Figure 7. Medium complexity

Figure 8. High complexity

Table 2. A quantitative assessment of the sight-reading materials[4)]

Complexity level												
mm.	**Low**				**Medium**				**High**			
1-4	10	8	8	10	8	7	11	8	8	10	12	10
5-8	8	9	9	8	9	10	12	8	13	12	10	12
9-12	6	10	7	7	11	10	11	11	14	13	18	13
13-16	8	6	9	7	12	11	11	10	11	10	12	12
Total notes	130				160				190			
Accidental	0				27				84			
Half note	19				8				0			
Key	C Major											
Meter	4/4											
Length	16 measures; 48 seconds											
BPM	80											

Examination of the Sight-Reading Materials

To obtain objective validity of the sight-reading materials, the degree of complexity was investigated. Specifically, the entropy of the sight-reading materials was calculated, and the value of entropy was compared with that of the greatest composers from different periods (Knopoff and Hutchinson, 1983; Youngblood, 1958). Entropy is derived from the information theory and is a mathematical tool for measuring an information value (Shannon and Weaver, 1949). Entropy H is determined by the following formula:

4) mm represents measure numbers and indicates the number of notes in each measure; total notes, accidentals, and half notes represent the number of total notes, accidentals, and notes with half notes or over half notes, respectively.

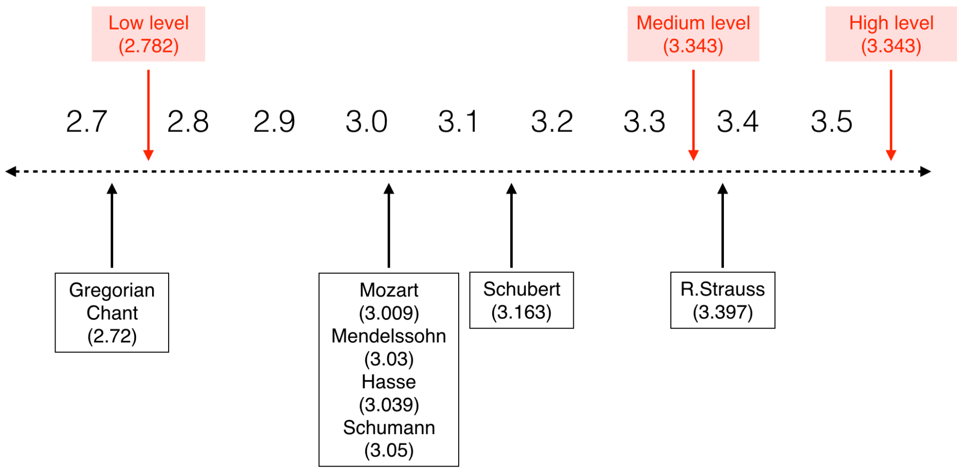
$$H(X) = -\sum_{i=1}^{n} (P_i \log_2 P_i)$$

Equation (1)

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

In music research, the number of probabilities is the number of pitch classes. There is a maximum of twelve probabilities. Entropy is higher when the twelve notes appear with equal frequency. The music of the twelve-tone technique in which all twelve notes are given equal importance is the most complex music. The entropy of music, therefore, represents the degree to which a piece is chromatic. To demonstrate the degree of chromaticism of the sight-reading materials, the entropy of the materials was calculated and compared to several references from Youngblood (1958) and Knopoff and Hutchinson (1983). Figure 9 shows a comparison of the entropies between the sight-reading materials and references with different styles of music. In Figure 9, entropy is shown to gradually increase throughout the history of Western music from Gregorian Chants in the Middle Ages to the 20th century. The simple piece was located between the Middle and Classical periods, and the complex piece has almost an equivalent entropy value to that of twelve-tone music. Through this comparison, it can be speculated how chromatic the sight-reading materials are in comparison to the works of Western classical composers.

Figure 9. Comparison of the degree of chromaticism between sight-reading materials and different styles of Western music[5)]



To objectively characterize the sight-reading materials related to note duration, the number of notes per second (NNS) was calculated and compared to those of references. The NNS represents the degree of note duration. Several existing pieces with three different types of tempo (i.e., slow, medium and high) were selected as reference pieces. Each piece selected as a reference for each tempo in the NNS analysis was generally regarded as a slow, mostly moderate or fast piece. J. S. Bach's "Air on the G String," the second movement of his orchestral suite No.3 in D Major BWV 1068, was selected for the slowest level. Franz Liszt's "Liebestraum No.3 in A flat Major S.541" was used for the medium level. "The Flight of the Bumblebee", an orchestral interlude from the opera "The tale of Tsar Saltan" composed by Nikolai Rimsky-Korsakov, was chosen for the fastest level. Since the BPM of the pieces varied according to recordings, the

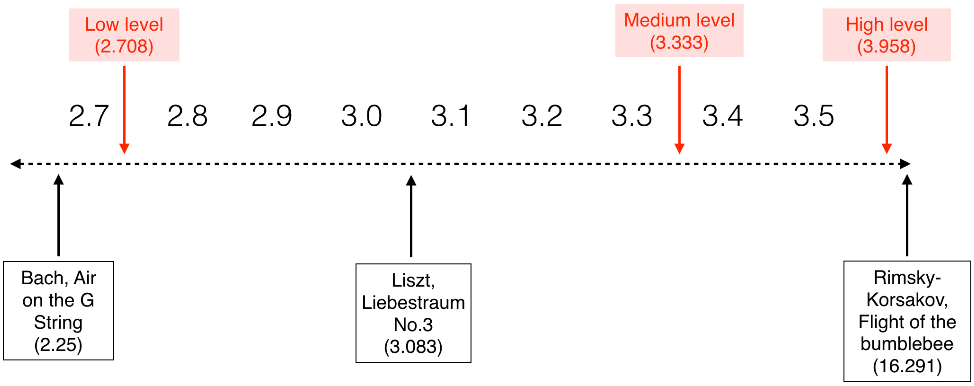
5) The bottom line of the middle-dashed arrow shows the entropy value of each composer based on the references. The upper end indicates the entropy value for sight-reading materials.

version that was watched the most on YouTube was chosen as a representative recording for each piece (see Table 3 for the selection of recordings). Figure 10 shows the NNS comparison between reference pieces and the sight-reading materials. The rhythmic complexity of the sight-reading materials was estimated.

Table 3. Selection of the reference pieces and recordings.

	Reference piece	
Slow tempo	J. S. Bach, "Air on the G String"	
	Recording performed by the Berlin Philharmonic Orchestra with Herbert von Karajan in 1966	27 BPM
Medium tempo	F. Liszt, "Liebestraum No. 3"	
	Recording performed by Yundi Li in 2003	100 BPM
Fast tempo	N. Rimsky-Korsakov, "The Flight of Bumblebee" (arr. by Sergei Rachmaninoff)	
	Recording performed by Evegeny Kissin in 2008	198 BPM

Figure 10. Comparison of the degree of note duration between the sight-reading materials and reference pieces.



3.2.3 Equipment

Musical pieces were written using the music notation software MuseScore 2.0.2. Participants sight-read and played the scores on a YAMAHA CLP-525 Clavinova digital piano, and their performances were directly recorded into Logic Pro X 10.2.2. in MIDI format.

3.2.4 Procedure

For the three musical pieces, the order of playing of the pieces was randomized and counterbalanced across participants. During the sight-reading performance, a metronome was provided for each beat to help participants maintain a constant tempo. Participants listened to the metronome beats for two measures before performing each score, and they were instructed to maintain a steady tempo as much as possible.

3.3 Evaluation

To evaluate the sight-reading proficiency of participants, both objective and subjective evaluations were used. Specifically, I adapted a dynamic time warping (DTW) algorithm for objective evaluations and calculated the distance between the reference and performance data. For subjective evaluations, I conducted a human rating experiment with ten experts and asked them to score the participants' performance. Details for each evaluation are described as follows.

3.3.1 Objective Evaluation

DTW is a method to identify an optimal cost path by tracking minimum costs within a similarity matrix of two sequences that vary in time and speed (Müller, 2007; Soulez et al., 2003). It uses an algorithm that differentiates the similarity between several consecutive data points and has been used effectively in the fields of speech and text recognition, as well as in score following. I used this computational method to provide a single value that would reflect the distance of the two different streams of music that should have been similar. The process of DTW can be divided into two sections: computing the similarity matrix and calculating the DTW distance. First, I applied the Jaccard distance to obtain a similarity matrix between each the performance of each participant and the sight-reading material. The Jaccard distance is widely used for the management of asymmetric binary variables (Lipkus, 1999; Willett and Barnard, 1998). The Jaccard distance between binary strings A and B can be calculated as following:

$$d_j(A,B) = 1 - J(A,B) = \frac{|A \cup B| - |A \cap B|}{|A \cup B|}$$

Equation (2)

where $J(A,B)$ is the Jaccard similarity coefficient. To compute this distance between each performance and the reference, each MIDI file was initially converted into binary matrices of values of 0 and 1. The first column of the performance matrix and the first column of the reference matrix used as inputs A and B of Equation (2), respectively. The remaining columns of each matrix were treated in the same manner, thereby yielding a group of Jaccard similarity coefficients. The Jaccard similarity coefficients of each matrix formed a single similarity matrix.

Subsequently, the DTW distance was computed from each similarity matrix, the process of which can be represented as the DTW algorithm.

This algorithm searches for a local minimum among the peripheral coordinates of each element of the matrix, including itself. When the local minimum is found, the value of the minimum in added to the element on the diagonal side. In other words, the values of the similarity matrix are then filled with new values, which ultimately form a new warping matrix. This process is shown in Figure 11, and can be summarized in Equation (3):

$$C(i,j) = \min \begin{cases} C(i,j-1) + D(i,j) \\ C(i-1,j) + D(i,j) \\ C(i-1,j-1) + D(i,j) \end{cases}$$

Equation (3)

where $C(i,j)$ is a coordinate of the warping matrix and is also the coordinate of the previous similarity matrix. The location of each local minimum indicates the direction in which the cost path should proceed within the matrix, starting from the zeropoint. This path represents the optimal path for matching the two matrices. Notably, the coordinates of the local minimum form the optimal cost path for the warping matrix, which is described in Figure 12. The value of the last coordinate of the optimal path is equal to the DTW distance. The distance for each performance was collected as raw data for the evaluation.

Figure 11. An example of the process for obtaining the warping matrix from the similarity matrix. The figure on the left indicates how to find the local minimum among a particular element (bolded border in the top-left corner) and its peripheral elements (dotted border) in the similarity matrix. In this figure, the value 1 is the local minimum. The figure on the right depicts how the local minimum is added to the element diagonal to the particular element (in the top left corner).

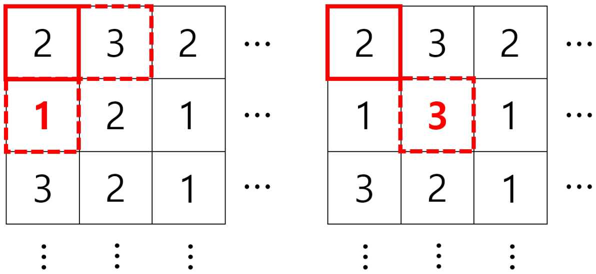
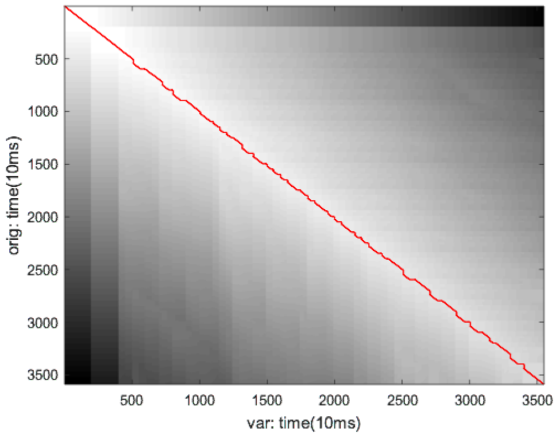


Figure 12. The warping matrix created from the DTW algorithm. The red line represents the optimal cost path of the warping matrix.



Using this algorithm, overall distances, including pitch and temporal information, between each of the performances and references were calculated.

3.3.2 Subjective Evaluation

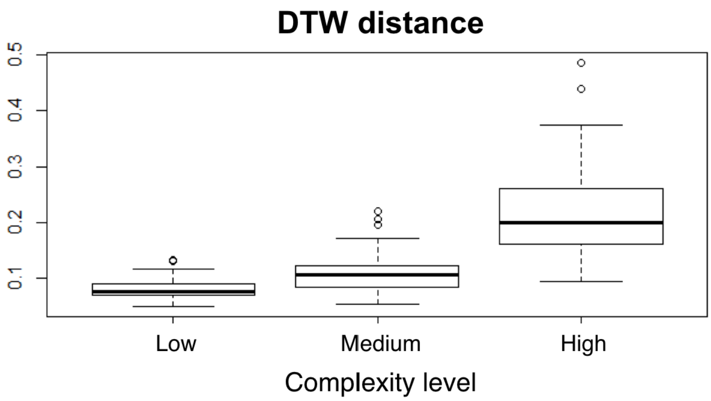
To test the reliability of the DTW algorithm, a human rating experiment was conducted as a subjective evaluation. The experiment was conducted as an off-line survey. A total of 10 professional pianists participated (9 females and 1 male; age range, 18-24 years; mean age=20.4 years) in the survey. A total of 92 performance data with three different complexities performed by 32 pianists from the sight-reading experiment were used in audio form in the human rating. Before starting to rate a score, participants were informed about the task and presented with references performed by a computer for each piece of varying complexity. During the experiment, participants listened to the audio clips for each level of complexity, including a 10-second pause between each performance, and were asked to rate the technical accuracy regarding pitch and duration on a scale from 1 to 5 (1=bad, 2=poor, 3=fair, 4=good and 5=excellent). The order of listening to each level of complexity was randomized and counterbalanced across participants. The subjective evaluation experiment took ~90 minutes to complete. After completing all of the evaluations, the results of the objective and subjective evaluations were compared with each other. Correlations between them were also examined in a cross-check.

3.4 Results

3.4.1 Objective Evaluation

One-way ANOVA was performed on data from the DTW distance to investigate the influence of musical complexity on sight-reading performance. The results are described in Figure 13. The box plot shows how the overall distance increases with the increasing complexity of music (low level: mean=0.0811; medium level: mean=0.1076; high level: mean=0.2134). Differences between the complexity conditions were significant ($p<.001$, MSE=0.314). This finding indicates that the complexity of music significantly influenced sight-reading performance by a somewhat large magnitude. Interestingly, the difference between the medium and high levels of complexity was even more greater than that between the low and medium levels. Additionally, the overall sample variance for each complexity level became larger with the increasing complexity of music ($p<.001$).

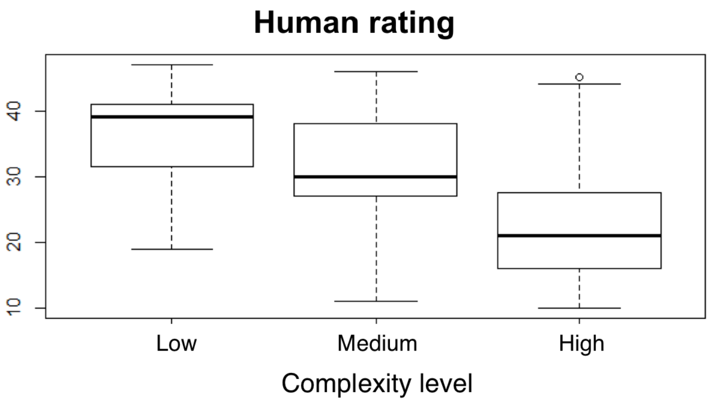
Figure 13. The overall distance with the increasing complexity of music. DTW distance is presented as a relative value that was normalized between 0 and 1.



3.4.2 Subjective Evaluation

One-way ANOVA was performed on data from the human rating score to investigate the influence of musical complexity on sight-reading performance. The results are shown in Figure 14. Similar to the results for the DTW distance, the box plot indicates how the human rating score decreases as the level of complexity increases (low level: mean=36.2031; medium level: mean=31.2187; high level: mean=22.2812). Additionally, significant differences among different complexity conditions were identified ($p<.001$), and the overall sample variance also became larger with the increasing complexity.

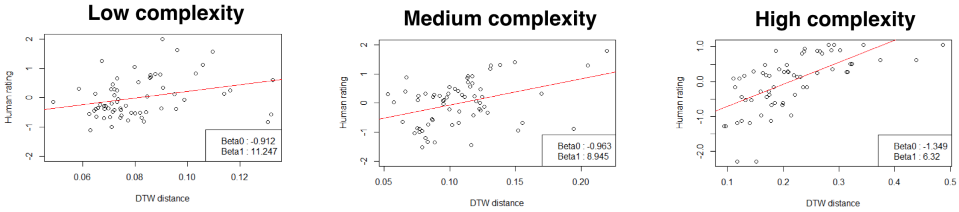
Figure 14. Changes in the human rating score with the increasing musical complexity.



3.4.3 Correlations between the Evaluations

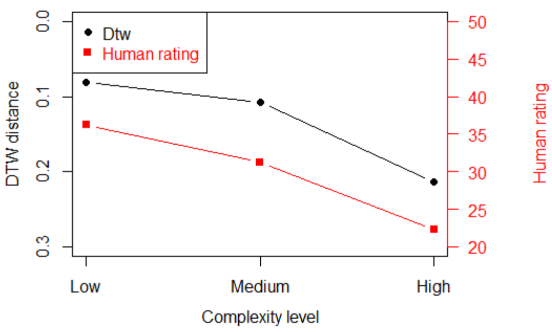
To test for a correlation between the objective and subjective evaluations, the distribution of data between the DTW distance and human rating value was compared. Since the DTW distance had a continuous range of values compared to the human rating score, which had a discrete value, the Kendall rank correlation coefficient was performed to test for a nonparametric relationship between the two measurements. The results shown in Figure 15 indicate a linear relationship between the two evaluations. It showed a highly significant agreement between the DTW distance and human rating score for all three levels of the complexity (low level: tau=.238, $p<.05$; medium level: tau=.36, $p<.05$; high level: tau=.69, $p<.001$). Interestingly, a positive correlation between the objective and subjective evaluations became greater as the level of the complexity increased.

Figure 15. A linear correlation between the DTW distance and human rating score for each level of complexity (Beta0: y-intercept; Beta1: slope)



Consequently, it can be concluded that the DTW algorithm is generally congruent with human perception, confirming the reliability of the DTW. Figure 16 shows the combined results of the DTW distance and human rating for the influence of musical complexity on sight-reading proficiency. The results show a lower tendency with increasing complexity for both evaluations conducted in parallel.

Figure 16. Combined results of the DTW distance and human rating for the influence of musical complexity on sight-reading proficiency. The left y-axis represents the inverse DTW distance and the right y-axis represents the human rating score.



3.5 Discussion

The study aimed to quantify musical complexity, taking into account chromatic and rhythmic components. The validity of the sight-reading material was examined by calculating the entropy and NNS. It also aimed to investigate the influence of musical complexity, which was newly proposed in this experiment, in an objective manner. To do this, I adapted the DTW algorithm for objective evaluation and examined the reliability of the algorithm through a human rating experiment. The experimental results showed that sight-reading proficiency significantly differed depending on the level of complexity in both the objective and subjective evaluations. This discussion focuses on two major points. First, participants' sight-reading proficiencies significantly differed according to musical complexity, and the sample variance became greater with an increasing complexity of music. Second, the steepness of the slope of the correlation between the objective and subjective evaluations also became increased as the music became more complex. The following section describes these factors in more detail along with their implications.

The Influence of Musical Complexity on Sight-Reading Accuracy

In both the DTW distance and human rating score, participants' sight-reading performances became less accurate as the complexity of music increased. In a sense, there is no doubt that performers would be proficient in sight-reading if a sight-reading piece was easy, and vice versa. Notable, however, is that although all of the participants were highly skilled and had a similar musical experience on average, they showed the different sight-reading accuracies between the different levels of complexity. Specifically, the sample variance increased according to the level musical complexity. This implies that individual differences, regardless of expertise, are the major determinant of sight-reading achievement when a sight-reading task has the highest level of complexity. Additionally, these results suggest a threshold for the level of the complexity in which practice-independent factors become the most important parameter (i.e., rather than practice-dependent factors) in sight-reading proficiency.

In this experiment, qualitative characteristics of music, such as pitch and note duration, were presented as criteria for musical complexity. The degree of chromaticism and note duration of the sight-reading materials was also demonstrated by calculating the entropy and number of notes per second of the pieces and by comparing those values with those of the references. To briefly summarize this process, the entropy value was located between the Gregorian chant in the Middle age and Mozart of the Classical period at low complexity, between the Romantic and late Romantic periods at medium complexity, and finally almost equivalent with Schoenberg's twelve-tone serial music at high complexity. In other words, there was little proficiency difference among the expert pianists when the sight-reading material had a similar degree of chromaticism to the Classical or Romantic period music. This indicates that the long-term consequences of deliberate practice are significant for the sight-reading of tonal music. In contrast, individual variations became greater when the sight-reading material had the highest degree of chromaticism, such as for Schoenberg's serialism music, which suggests that the importance of practice-independent factors, such as general and elementary cognitive skills, on sight-reading accuracy. This assumption is supported by several other studies. Platz et al. (2014) and Ericsson et al. (1993) examined the significance of deliberate practice for outstanding sight-reading performances. Ericsson and Lehmann (1996) concluded that the best sight-readers were accompanists who were trained to perform with limited or no preparation. Thus, it can be assumed that the higher accuracy of participants' sight-reading performance in the low and medium complexity levels, which represent tonal music, can be attributed to practice-dependent factors such as deliberate practice or the size of sight-reading repertoires. Rather than the significance of expertise, the importance of practice-independent factors, such as innate cognitive abilities, on superior sight-reading performance were also demonstrated. For example, working memory capacity (Meinz and Hambrick, 2010; Vandervert, 2009), handedness (Kopiez et al., 2010; 2006), sensorimotor speed (Kopiez and Lee, 2008; 2006), intrinsic motivation (Winner, 1996), unique types of representations (Shavinina, 2009) and verbal memory (Brandler and Rammsayer, 2003) were examined as practice-independent factors. Therefore, the large sample variance at the high level of complexity corresponding to atonal music among the expert pianists could be explained by practice-independent factors.

Taken together, one interpretation of the results of this experiment is that factors that influence sight-reading accuracy can vary with musical complexity. This is congruent with the findings of Kopiez and Lee (2008; 2006). They investigated and ranked the predictors of sight-reading achievement, and found that the combination of predictor variables changed with the complexity of music. Notably, with increasing sight-reading task difficulty, the importance of psychomotor speed, speed of information processing, inner hearing, and sight-reading expertise becomes greater. When sight-reading tasks reach the highest complexity level of music, sight-reading expertise remains important, but psychomotor speed, such as trilling speed, becomes the dominant predictor. However, as only quantitative characteristics of music, such as the total number of notes or the total duration of a piece, were considered in the difficulty of the sight-reading material and all of the materials were tonal music, it was difficult to identify which qualitative characteristics of music were important to musical complexity. Based on the findings of Kopiez and Lee (2008; 2006), this experiment first examined whether the same results could be found under different complexity conditions, considering qualitative properties of music such as chromatic and note duration components. Consequently, we found a conversion of a combination of sight-reading predictors between the medium and high complexity levels, among which the degree of chromaticism matched with R. Strauss and twelve-tone music, respectively.

Correlations between the Objective and Subjective Evaluations

Another interesting point about these experimental results is that the correlation coefficient between the objective and subjective evaluations also differed according to musical complexity: the slope of the correlations between the DTW distance and the human rating score continued to become steeper and steeper towards a value of ‘1’ as the level of complexity increased. Similarities between the two evaluations were directly proportional to increasing complexity, and such analogous scoring between the evaluations implies that human ability to judge sight-reading accuracy improves as a musical piece becomes more difficult. Because the sight-reading materials used in this experiment had no dynamic or articulation directions, and only the pitch and duration difficulties were considered in the division of the complexity level. The most important criterion in the evaluation of the participants’ sight-reading performances was only whether they performed the score accurately or not. Consequently, the strong consensus between the objective and subjective evaluations demonstrates the reliability of the DTW adaptation to sight-reading evaluation as well as the improvement of humans’ perceptual judgement abilities in accordance with increasing musical complexity.

Chapter 4. Physiological Experiment

4.1 Aims

Many previous studies have argued that the length of the EHS can vary according to sight-reading skill levels (Cara, 2017; Rosemann et al., 2015; Gilman and Underwood, 2003; Furneaux & Land, 1999; Truitt et al., 1997; Sloboda, 1974). These studies suggested that the quality of performance, such as performance accuracy or playing tempo, reflects the differential value of the EHS. For example, sight-readers who performed more accurately (Cara, 2017; Rosemann et al., 2015; Gilman and Underwood, 2003; Sloboda, 1974) or faster (Cara, 2017; Furneaux and Land, 1999; Truitt et al., 1997) have been regarded as having a larger EHS. Among them, however, only a few studies have directly examined correlations between the EHS and quality of performance (Rosemann et al., 2015; Gilman & Underwood, 2003; Sloboda, 1974). Interestingly, Sloboda (1974) first demonstrated a strong positive correlation between sight-reading skills and performance accuracy, but a recent study showed only a slight trend for a correlation between these two variables (Rosemann et al., 2015). In other words, the length of the EHS did not correspond to the quality of performance. This fact enables the assumption that professional pianists may not differ in looking further ahead and storing musical notations for sight-reading achievement. Based on these limitation, this study aimed to investigate correlations between the EHS and sight-reading performance. There was a distortion in the EHS results because the playing tempo varied according to the participants. Therefore, it was necessary to investigate whether the size of the EHS is proportional to sight-reading achievement in terms of performance accuracy rather than playing tempo.

Secondly, the EHS was commonly estimated to be approximately 1 second in several previous studies (Wurtz et al., 2009; Furneaux and Land, 1999). The EHS did not change according to the level of expertise (Furneaux and Land, 1999) or complexity of music (Wurtz et al., 2009). Only the playing tempo influenced the time index of the EHS. This implies that every sight-reader may read the score for a certain amount of time irrespective of how many notes or beats enter in a fixed time window. If so, a long-lasting debate on the EHS, such as how much farther and faster skilled pianists read musical notes than the less skilled pianists, would be pointless. The EHS cannot be an indicator of differences in sight-reading ability. Instead, other cognitive abilities, such as the speed of information processing, would be the dominant factor to explain differences in the EHS. Thus, it remains to be investigated whether the EHS is consistent with a note, beat, or absolute time. The reason for the disagreement among the EHS indexes is that playing tempo was uncontrolled in almost all of the studies. For example, each participant played at their own tempo, and performers who played faster were regarded as skilled sight-readers. The length of the EHS should be varied. Consequently, the EHS and relevant points should be examined under the controlled condition of playing tempo. This experiment, therefore, aimed to explore whether the EHS follows a note, beat or absolute time. Although the EHS in time index showed the consistency that was independent of any variable except playing tempo, there has been no study that simultaneously calculated the EHS in note, beat and time. Thus, the most valid unit to measure the EHS also needed to be verified.

Finally, the last goal of this study was to re-examine the influence of musical complexity and playing tempo on the EHS. Musical complexity, such as playing tempo, was also quantitatively uncontrolled in the previous studies. No standard or definition of the complexity was presented, so it is hard to speculate how much complexity influences the length of the EHS. I already suggested the degree of musical complexity as a quantitative metric in Chapter 3. By using the proposed level of complexity from the previous experiment, the study presented in Chapter 4 re-examines the influence of musical complexity on the EHS. Similarly, the influence of playing tempo on the EHS will be re-examined by fixing the playing tempo among participants.

4.2 Methods

4.2.1 Participants

A total of 35 students (34 females and 1 male; mean age=21.89 years, SD=2.040 years), who major in piano at Seoul National University, took part in this study. The participants had started piano playing at an average of 6.028 years (SD=1.014 years, range 4-7 years), and had played the piano for 15.857 years (SD=2.211 years, range 12-20 years) on average. Regarding professional experience, they had started to major in piano as their primary instrument at an average of 11.428 years (SD=2.329 years, range 6-16 years), and had majored in piano for 10.457 years (SD=2.893 years, range 5-18 years). Based on the accumulated time they had spent playing the piano expertly, I judged that they were all should be considered to be professional pianists in terms of general pianistic skills and that there were no novices or intermediate pianists among them. All of the participants were right-handed and had normal or corrected-to-normal vision.

4.2.2 Materials

In this physiological experiment, sight-reading materials were chosen from the experiment described in Chapter 3 because those pieces were composed under quantitative controls. Originally, out of music of three different complexities (low, medium and high), the medium level of complexity was excluded to maximize comparisons of the influence of different complexities. Therefore, four musical pieces with two different complexities (simple and complex) and tempi (slow and fast) were used in this experiment. Because there were two contrasting tempi for both complexities, two pieces out of four sight-reading materials were newly composed in the same way as for the experiment described in Chapter 3. Each complexity had two different versions of a piece, which only differed in tempo and consisted of quantitatively the same musical components with each other. All of the pieces had the same kay (C Major), time signature (4/4 meter) and length (16 measures) to prevent the introduction of confounding variables apart from the complexity and playing tempo.

4.2.3 Equipment

Sight-reading materials were presented on a 23′′ monitor with a resolution of 1920 × 1080 pixels. Binocular movements were recorded using Tobii Pro Glasses 2 (Stockholm, Sweden) with a sampling rate of 50 Hz (every 20 millisecond). The distance between the eye and monitor was 50 cm. Participants were required to stabilize their head as much as possible, but also were able to glance at their fingers to maintain a natural condition during sight-reading. A Yamaha CLP-525 Clavinova digital piano was used in the experiment, and the participants' performance was directly recorded into Logic Pro X 10.2.2. in MIDI format.

4.2.4 Procedure

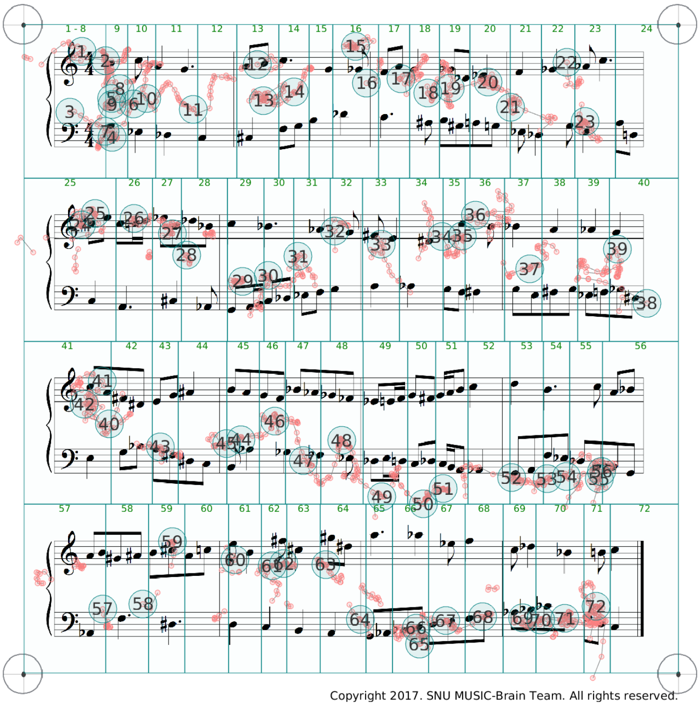
All of the participants played four sight-reading pieces (simple, slow; simple, fast; complex, slow; complex, fast) in a random order. Before starting each session, the participants calibrated their eye tracker at four different points on the sheet music. After the calibration phase, a metronome was provided for two measures before playing, and then the participants started to sight-read along with the metronome. The metronome was provided for each beat to help the participants maintain a constant tempo. In total, the experiment lasted approximately 20 minutes. After the main experiment, participants were asked to complete a short questionnaire about their musical experience.

4.3 Data Analysis

Eye Movements

To calculate the EHS, the eye tracking program was modified to simultaneously record the eye movements and receive MIDI input for the played notes. Temporal information about eye movements and onsets of note performance were directly acquired and plotted on sheet music projected on a monitor using Matlab-based programs. With this visualization, computation of the EHS was possible as both temporal information about eye movements and note performance as well as information about the sheet music itself was available. The EHS was calculated as 1) latency (in ms) between fixating and playing a note, 2) the number of beats and 3) the notes between the current eye and hand position. The EHS was computed for every beat of every measure, ultimately yielding 48 data points. The beat and time spans were proportional to each other because the time span was the beat span multiplied by playing tempo.

Figure 17. Visualization of the EHS-numbers indicate the EHS order represented as beat index.



Performance Accuracy

Sight-reading performance was evaluated by calculating the distance between reference and performance data using a dynamic time warping algorithm as in the previous experiment.

4.4 Results

4.4.1 Correlations between the EHS and Performance Accuracy

A correlation analysis was conducted to identify associations between the quality of performance (DTW distance) and length of the EHS (i.e., note, beat and time). No significant correlations were detected between the DTW distance and EHS in beat ($r(20)$=0.0005, p=0.9968) or time ($r(2)$=0.0703, p=0.5356). Instead, there was a slight negative correlation between the DTW distance and EHS in note ($r(20)$=−0.2359, p=0.0352). This indicates that a shorter EHS in note is associated with a more accurate sight-reading performance.

Table 4. Correlation between the DTW distance and the EHS

EHS	ρ	p-value
in beat	-0.1902	0.9968
in time	-0.2208	0.5356
in note	0.1102	0.0352

alpha $\alpha = 0.01$

4.4.2 Variable Factors

A linear mixed model was used to test the main effects and interactions of musical complexity and playing tempo on the EHS. In the simple piece, the average EHS was 3.47 in note, 1.68 in beat, and 1117 ms in time. In the complex piece, the avaerage EHS was 4.17 in note, 1.38 in beat, and 889 ms in time. The main effect of musical complexity was significant on the EHS in beat and time (p=0.02112 for beat; 0.0045461 for time). No main effect of complexity was found in the EHS for note (p=0.6953).

In the slow tempo condition, the mean EHS was 3.31 in note, 1.38 in beat, and 1032 ms in time. In the fast tempo condition, the average EHS was 4.32 in note, 1.69 in beat, and 974 ms in time. Playing tempo had no significant major effect across all of the EHS indexes (p=0.9925 for note; 0.99334 for beat; 0.068394 for time).

Table 5. Average EHS in note, beat, and time along with the complexity and playing tempo variables.

EHS	Complexity	Playing tempo
in beat	F=5.5447	F=7.0072x10^{-5}
	p=0.02112*	p=0.99334
in time	F=8.5523	F=3.4176
	p=0.0045461**	p=0.068394
in note	F=0.15648	F=8.8952x10^{-5}
	p=0.69353	p=0.9925

F=$F(1,76)$ *p<.05; **p<.01

4.4.3 EHS Examination

Based on the results shown in section 4.4.2, it was found that the EHS followed a consistent number of notes. It did not change according to musical variables. The length of the EHS in note of simple and complex pieces was similar, even though 30 more notes were in the complex piece than in the simple piece. Additionally, in note the EHS of slow and fast tempo conditions was similar, although the playing tempo was 30 percent faster in the fast tempo conditions. This means that whether the piece is simple or complex, and whether the playing tempo is slow or fast, professional pianists are looking for a certain number of notes more than at their hands. Therefore, in this study, the most valid unit of the EHS is the note.

4.5 Discussion

This experiment aimed to examine a correlation between the EHS and sight-reading proficiency. It also investigated the influence of variable factors on the length of the EHS and attempted to identify relationships between the various indexes describing the EHS. In summary, there were two major findings based on the experimental results. First, there was no correlation between sight-reading proficiency and length of the EHS. Secondly, the EHS followed a consistent number of notes regardless of performers' sight-reading achievement.

Correlations between the EHS and Performance Accuracy

Many previous studies have argued that the length of the EHS varied according to sight-reading skill levels (Cara, 2017; Rosemann et al., 2015; Gilman and Underwood, 2003; Furneaux & Land, 1999; Truitt et al., 1997; Sloboda, 1974). They suggested that the quality of performance, such as performance accuracy or playing tempo, reflects the differential value of the EHS. For example, sight-readers who performed more accurately (Cara, 2017; Rosemann et al., 2015; Gilman and Underwood, 2003; Sloboda, 1974) or faster (Cara, 2017; Furneaux and Land, 1999; Truitt et al., 1997) have been regarded as having a larger EHS. Among them, however, only a few studies have directly examined correlations between the EHS and the quality of performance (Rosemann et al., 2015; Gilman & Underwood, 2003; Sloboda, 1974). Interestingly, Sloboda (1974) was the first to demonstra a strong positive correlation between sight-reading skills and performance accuracy, but a recent study showed only a slight trend for a correlation between the two variables (Rosemann et al., 2015). The results of this experiment support those of the latter one. The reason for the opposing correlation results among the EHS studies might be because the skill level of the participants for sight-reading differed in each study. The subjects of Sloboda (1974) possessed a wide range of sight-reading ability. It seems obvious that the value of the EHS would vary as a consequence. Furthermore, as Gilman and Underwood (2003) pointed out, the technique used in Sloboda (1974) was not suitable for measuring the EHS. He evaluated participants' sight-reading ability by calculating the number of errors made prior to the slide being switched off. The EHS was measured by calculating the number of notes after the slide was turned off. Since this technique corresponds to measurements of both how much music notation pianists can predict and how much notation pianists can see, Sloboda (1974) was unable to distinguish between the EHS, note identification span, and perceptual span. In contrast, only professional pianists participated in the study by Rosemann et al. (2015) and the present study. Considering the consistent results that there was nearly no correlation between the EHS and sight-reading achievement, in the study by Rosemann et al. (2015) and the present study, it can be concluded that sight-reading proficiency and the EHS have no significant relationship, at least in case of professional pianists. This finding implies that working memory, which is represented by the EHS, may not be the critical difference that can explain sight-reading achievement. Instead, other cognitive abilities, such as the speed of information processing or psychomotor speed, may be dominant factors that could be used to clarify differences in sight-reading.

The Note-Consistent EHS

The EHS in time was commonly ~1 second in several previous studies (Wurtz et al., 2009; Furneaux and Land, 1999). Additionally, both objective and subjective variables, such as musical complexity or expertise, did not affect this time value. Only playing tempo influenced the time index of the

EHS. However, each participant played at their own tempo during the sight-reading task in these two studies. This experiment aimed to investigated whether a consistent lag between the hand and eye exists in a fixed playing tempo condition. Moreover, it also aimed to identify a valid unit to illustrate the EHS among the three values of absolute time, beat and note. No previous study has yet reported measurements of the EHS in all indexes simultaneously.

Surprisingly, the experimental results revealed that professional pianists looked at a certain number of notes further than their hands during sight-reading. The EHS in time and beat varied according to various musical variables, such as the complexity of music or playing tempo, but the EHS in note consistently held a similar value regardless of any variable in all of the experimental conditions that were assessed. This finding suggests that the participants steadily read a constant number of notes whether the sight-reading material was simple or complex. Considering the absolute number of notes used in the complex piece condition was greater than in the simple piece condition, such results are highly interesting in terms of identifying the number of musical units available for information processing.

It is generally known that the sight-reading strategy of skilled sight-readers is pattern recognition. However, the consistency of the EHS in note suggests that a meticulous reading in note units surpasses the chunking strategy. The participants in this experiment were professional pianists whose sight-reading skills were estimated to be high on average. The sight-reading strategy of participants was ascertained by eye tracking while paying scrupulous attention to their reading of music. Hence, the sight-reading strategy that professional pianists use for their competent performance at the input level is precisely reading a musical note rather than recognizing the notes to be merged into a larger unit.

Chapter 5. Conclusion

5.1 Summary and Implications of Main Findings

In sum, this study attempted to quantify musical complexity, taking into account the qualitative characteristics of music and investigating its influence on sight-reading proficiency. Pitch and durational difficulties were proposed as the criteria for musical complexity and each factor represented the degree of chromaticism and note duration, respectively. Additionally, the validity of the sight-reading material was examined by calculating and comparing the entropy and NNS between the sight-reading materials and references. This study adapted the DTW algorithm for objective evaluation and examined its reliability in human rating experiments. The results showed that sight-reading proficiency significantly differed depending upon the level of complexity in both objective and subjective evaluations. The accuracy of sight-reading performance decreased as the level of musical complexity increased, and the difference in sight-reading proficiency between the medium and high complexity levels was greater than that of the low and medium complexities. Interestingly, the variance became greater as the complexity of music increased, even though all the participants were expert pianists. Therefore, this study suggests that practice-independent factors are the most prominent determinant of sight-reading proficiency when the sight-reading task reaches the highest level of complexity.

Significant linear correlations between the objective and subjective evaluations were found for all complexity levels, and the correlation was positively greater with an increasing level of complexity. This finding implies that discriminatory abilities related to sight-reading evaluation improve as sight-reading materials become more difficult. Moreover, the significant congruence between the objective and subjective evaluations at a high level of complexity demonstrates the substantial reliability of the DTW algorithm for sight-reading evaluations. Overall, these findings suggest that sight-reading proficiency significantly differs depending upon musical complexity, particularly when the degree of chromaticism reaches a highest level of music, such as twelve-tone music. The predictor variables of sight-reading proficiency change with an increasing or decreasing degree of chromaticism or note duration of sight-reading materials. Additionally, the perceptual judgement abilities of humans for sight-reading accuracy improve as the level of complexity increases.

This study also attempted to resolve and complement the limitations of previous studies related to the quantification of sight-reading materials and playing tempo. Accordingly, this study examined the essential nature of the EHS, which has been considered as an indicator that represents differences in sight-reading proficiency. This study challenged the premise that the length of the EHS is proportional to sight-reading skills and scrutinized the correlation between these factors under highly restricted and organized experimental conditions. Its findings demonstrated the EHS was irrelevant to performance accuracy. This outcome contradicts a series of previous studies that investigated the influence of various musical variables on the EHS based on the major proposition that the EHS reflects variable sight-reading abilities. In other words, cognitive processing skills prior to motor output from a sensory input are likely to not be associated with proficient sight-reading performances. If so, at what stage of brain processes is related to differences in sight-reading skills? A discourse on the EHS is needed to understand this issue.

Irrespective of irrelevancy between the EHS and the accuracy of sight-reading performance, this study identified the way in which professional pianists input and process visual information during sight-reading. Its findings indicated that their sight-reading strategy was reading music meticulously (i.e., constantly reading a certain number of notes). A prevailing theory has understood the sight-reading strategy of skilled sight-readers as grouping individual notes into a larger unit and recognizing the musical patterns. In the previous studies, however, the level of the participants was not as high as in this present study, and the experiments were conducted under the unsystematic experimental conditions, such as an unfixed playing tempo. As a result, generalizations of the past experimental results are likely to be confounded. Instead, this present study suggests that dedicated music reading is a high level sight-reading strategy and proposes the need for a novel and dynamic model to explain differences in sight-reading proficiency.

This study began with a theoretical review of the neurological and physiological mechanisms of sight-reading, and then explored the influence of musical characteristics operating as variable elements in terms of complexity and playing tempo in an experiential and scientific manner. If sight-reading abilities are one of the most important aptitudes for becoming an excellent performer, the series of studies on sight-reading performance included in this study will represent a chance to present a fundamental definition of musical ability. On the other hand, many studies on differences in sight-reading ability have been fundamentally concerned with improvements in sight-reading skills. It is hoped that future theoretical and empirical studies of sight-reading will not only end with an exploration of the mechanisms and variables related to sight-reading performance, but will also lead to a new approach for effective practice designed to improve sight-reading performance.

5.2 Limitations and Further Research

Only professional pianists participated in this study. Thus, it is difficult to generalize these findings about sight-reading skills to a wider range of sight-readers. Studies of less skilled or non-professional pianists and comparisons of those results with those of professionals should be the subject of future studies. This would help to lead a better understanding of underestimated matters related to the EHS.

Bibliography

Brandler, S., & Rammsayer, T. H. (2003). Differences in mental abilities between musicians and non-musicians. *Psychology of Music*, *31*(2), 123-138.

Burman, D. D., & Booth, J. R. (2009). Music rehearsal increases the perceptual span for notation. *Music Perception: An Interdisciplinary Journal*, *26*(4), 303-320.

Cara, M. A. (2017). Anticipation awareness and visual monitoring in reading contemporary music. *Musicae Scientiae,* 1029864916687601.

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.

Ericsson, K. A., Krampe, R., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(39), 363-406.

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology*, *47*(1), 273-305.

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

Gilman, E., & Underwood, G. (2003). Restricting the field of view to investigate the perceptual spans of pianists. *Visual Cognition, 10*(2), 201-232.

Goolsby, T. W. (1994a). Eye movement in music reading: Effects of reading ability, notational complexity, and encounters. *Music Perception: An Interdisciplinary Journal, 12*(1), 77-96.

Goolsby, T. W. (1994b). Profiles of processing: Eye movements during sightreading. *Music Perception: An Interdisciplinary Journal*, *12*(1), 97-123.

Jäncke, L., Kleinschmidt, A., Mirzazade, S., Shah, N. J., & Freund, H. J. (2001). The role of the inferior parietal cortex in linking the tactile perception and manual constructon of object shapes. *Cerebral Cortex, 11*(2), 114-121.

Kinsler, V., & Carpenter, R. H. S. (1995). Saccadic eye movements while reading music. *Vision Research, 35*(10), 1447-1458.

Knopoff, L., & Hutchinson, W. (1981). Information theory for musical continua. *Journal of Music Theory, 25*(1), 17-44.

Kopiez, R., & In Lee, J. (2006). Towards a dynamic model of skills involved in sight-reading music. *Music Education Research*, *8*(1). 97-120.

Kopiez, R., & In Lee, J. (2008). Towards a general model of skills involved insight-reading music. *Music Education Research*, *10*(1). 41-62.

Kopiez, R., Galley, N., & Lee, J. I. (2006). The advantage of a decreasing right-hand superiority: The influence of laterality on a selected musical skill (sight reading achievement). *Neuropsychologia*, *44*(7), 1079-1087.

Kopiez, R., Galley, N., & Lehmann, A. C. (2010). The relation between lateralization, early start of training, and amount of practice in musicians: a contribution to the problem of handedness classification. *Laterality*, *15*(4), 385-414.

Lehmann, A. C., & Ericsson, K. A. (1996). Performance without preparation: structure and acquisition of expert sight-reading and accompanying performance. *Psychomusicology*, *15*(1/2), 1-29.

Lehmann, A. C., & Kopiez, R. Sight-reading. In S. Hallam, I. Cross & M. Thaut (Eds.) (2009). *The Oxford Handbook of Music Psychology* (pp. 344-351). Oxford: Oxford University Press.

Lipkus, A. H. (1999). A proof of the triangle inequality for the tanimoto distance, *Journal of Mathematical Chemistry*, *26*, 263-265.

Madell, J., &Hééebert, S. (2008). Eye movements and music reading: Where do we look next?. *Music Perception: An Interdisciplinary Journal*, *26*(2), 157-170.

McPherson, G. E. (1993). *Factors and abilities influencing the development of visual, aural, and creative performance skills in music and their educational implications* (Doctoral dissertation, University of Sydney).

McPherson, G. E. (1995). Five aspects of musical performance and their correlates. *Bulletin of the Council for Research in Music Education*, *127*, 115-121.

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), 914-9.

Müller, M. (2007). Dynamic time warping. *Information Retrieval for Music and Motion* (pp.69-84). Heidelberg: Springer.

Penttinen, M. (2013). *Skill development in music reading: The eye-movement approach* (Doctoral dissertation, University of Turku). Retrieved from http://www.doria.fi/handle/10024/88757.

Platz, F., Kopiez, R., Lehmann, A. C., & Wolf, A. (2007). The influence of deliberate practice on musical achievement: a meta-analysis. *Psychological Perspectives on Expertise*, *151*.

Podolak, O. M. (2013). *The influence of tonality on sight-reading accuracy* (Doctoral dissertation, University of Toronto).

Pollatsek, A., Rayner, K., & Balota, D. A. (1986). Inferences about eye movement control from the perceptual span in reading. *Perception &*

Psychophysics, *40*(2), 123-130.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372-422.

Rayner, K. (1993). Eye movements in reading: Recent developments. *Current Directions in Psychological Science, 2*(3), 81-86.

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.

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.

Sergent, J., Zuck, E., Terriah, S., & MacDonald, B. (1992). Distributed neural network underlying musical sight-reading and keyboard performance. *Science*, *257*(5066), 106-109.

Schön, D., Anton, J. L., Roth, M., & Besson, M. (2002). An fMRI study of music sight-reading. *Neuroreport*, *13*(17), 2285-2289.

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

Shavinina, L. V. (2009). A unique type of representation is the essence of giftedness: Towards a cognitive-developmental theory. In *International Handbook on Giftedness* (pp. 231-257). Springer, Dordrecht.

Simon, O., Mangin, J. F., Cohen, L., Le Bihan, D., & Dehaene, S. (2002). Topographical layout of hand, eye, calculation, and language-related areas in the human parietal lobe. *Neuron, 33*(3), 475-487.

Sloboda, J. (1974). The eye-hand span-An approach to the study of sight reading. *Psychology of Music, 2*(2), 4-10.

Sloboda, J. (1984). Experimental studies of music reading; A review. *Music*

Perception, 2(2), 222-236.

Soulez, F.,Rodet, X. & Schwarz, D. (2003). Improving polyphonic and poly-instrumental music to score alignment. *Proceedings of International Society on Music Information Retreival,* 6-11.

Stewart, L. (2005). A neurocognitive approach to music reading. *Annals of the New York Academy of Sciences, 1060*(1), 377-386.

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.

Vandervert, L. R. (2009). Working memory, the cognitive functions of the cerebellum and the child prodigy. In *International Handbook on Giftedness* (pp.295-316). Springer, Dordrecht.

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

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

Willett, P., Barnard, J. M. & Downs, G. M. (1998). Chemical similarity searching, *Journal of Chemical Information and Computer Sciences, 38*, 983-996.

Winner, E., & Drake, J. E. (1996). The rage to master: The decisive role of talent in the visual arts. *Kaufman, SB: The Complexity of Greatness: Beyond Talent or Practice*, 333-366.

Wolf, T. (1976). A cognitive model of musical sight-reading. *Journal of Psycholinguistic Research, 5*(2), 143-171.

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.

Youngblood, J. E. (1958). Style as information. *Journal of Music Theory*, *2*(1). 24-35.

국문 초록

눈-귀-손 협력 과정으로서의 초견 연주 인지 모델

음악 연주는 행동 및 인지의 분류, 그리고 그것의 생리학적 기반에 대한 연구가 필요한 실제 세계의 활동을 나타낸다. 본 논문은 기본적이지만 중요한 음악적 능력인 초견을 다룬다. 초견이란 고도로 복잡한 시각, 청각, 운동 영역의 처리 작업을 포함하고 있기 때문에, 초견의 인지 메커니즘에 대한 연구는 음악 분야 안밖에서 인간 행동과 인지를 근본적으로 이해하는데 기여할 것이다. 본 논문은 이론적인 차원에서 초견의 신경학적 생리학적 메커니즘을 탐구하고, 초견의 메커니즘에 영향을 미치는 다양한 요인을 경험적 증거에 기반하여 탐구하는 것을 목표로 한다. 특히 본 논문은 선행 연구에서 부족했던 기술적 한계와 음악적 한계를 지적하고 보완한다. 궁극적으로 본 논문은 초견에 대한 새로운 인지 모델을 제시한다.

주요어 : 초견, 시선추적, 눈-손 간격, 복잡성, 엔트로피, DTW

학번 : 2016-21879