TURUN YLIOPISTON JULKAISUJA
ANNALES UNIVERSITATIS TURKUENSIS

SARJA - SER. B OSA - TOM. 359

HUMANIORA

SKILL DEVELOPMENT IN MUSIC READING

The Eye-Movement Approach

by

Marjaana Penttinen

TURUN YLIOPISTO
UNIVERSITY OF TURKU
Turku 2013

From the Department of Teacher Education and the Centre for Learning Research
University of Turku, Finland

Supervisors:

Professor Mirjamaija Mikkilä-Erdmann
Department of Teacher Education
Centre for Learning Research
University of Turku, Finland

Professor Erno Lehtinen
Department of Teacher Education
Centre for Learning Research
University of Turku, Finland

Professor Emeritus Erkki Olkinuora
Department of Teacher Education
Centre for Learning Research
University of Turku, Finland

Reviewers and opponents:

Professor Hans Gruber
Institute for Educational Science
University of Regensburg, Germany

Dr. Halszka Jarodzka
Centre of Learning Sciences and Technologies
Open Universiteit Nederland, the Netherlands

Cover: “Elektroninen valo” by Odin Puurtinen

ISBN 978-951-29-5329-5 (PRINT)
ISBN 978-951-29-5330-1 (PDF)
ISSN 0082-6987
Painosalama Oy – Turku, Finland 2013

ABSTRACT

This dissertation examined skill development in music reading by focusing on the visual processing of music notation in different music-reading tasks. Each of the three experiments of this dissertation addressed one of the three types of music reading: (i) sight-reading, i.e. reading and performing completely unknown music, (ii) rehearsed reading, during which the performer is already familiar with the music being played, and (iii) silent reading with no performance requirements. The use of the eye-tracking methodology allowed the recording of the readers' eye movements from the time of music reading with extreme precision. Due to the lack of coherence in the smallish amount of prior studies on eye movements in music reading, the dissertation also had a heavy methodological emphasis. The present dissertation thus aimed to promote two major issues: (1) it investigated the eye-movement indicators of skill and skill development in sight-reading, rehearsed reading and silent reading, and (2) developed and tested suitable methods that can be used by future studies on the topic.

Experiment I focused on the eye-movement behaviour of adults during their first steps of learning to read music notation. The longitudinal experiment spanned a nine-month long music-training period, during which 49 participants (university students taking part in a compulsory music course) sight-read and performed a series of simple melodies in three measurement sessions. Participants with no musical background were entitled as "novices", whereas "amateurs" had had musical training prior to the experiment. The main issue of interest was the changes in the novices' eye movements and performances across the measurements while the amateurs offered a point of reference for the assessment of the novices' development. The experiment showed that the novices tended to sight-read in a more stepwise fashion than the amateurs, the latter group manifesting more back-and-forth eye movements. The novices' skill development was reflected by the faster identification of note symbols involved in larger melodic intervals. Across the measurements, the novices also began to show sensitivity to the melodies' metrical structure, which the amateurs demonstrated from the very beginning. The stimulus melodies consisted of quarter notes, making the effects of meter and larger melodic intervals distinguishable from effects caused by, say, different rhythmic patterns.

Experiment II explored the eye movements of 40 experienced musicians (music education students and music performance students) during temporally controlled rehearsed reading. This cross-sectional experiment focused on the eye-movement effects of one-bar-long melodic alterations placed within a familiar melody. The synchronizing of the performance and eye-movement recordings enabled the investigation of the eye-hand span, i.e., the temporal gap between a performed note and the point of gaze. The eye-hand span was typically found to remain around one second. Music performance students demonstrated increased professing efficiency by their shorter average fixation durations as well as in the two examined eye-hand span measures: these participants used larger eye-hand spans more frequently and inspected more of the musical score during the performance of one metrical beat than students of music education. Although all participants produced performances almost indistinguishable in terms of their auditory characteristics, the altered bars indeed affected the reading of the score: the general effects of expertise in terms of the two eye-hand span measures, demonstrated by the music performance students, disappeared in the face of the melodic alterations.

Experiment III was a longitudinal experiment designed to examine the differences between adult novice and amateur musicians' silent reading of music notation, as well as the changes the 49 participants manifested during a nine-month long music course. From a methodological perspective, an opening to research on eye movements in music reading was the inclusion of a verbal protocol in the research design: after viewing the musical image, the readers were asked to describe what they had seen. A two-way categorization for verbal descriptions was developed in order to assess the quality of extracted musical information. More extensive musical background was related to shorter average fixation duration, more linear scanning of the musical image, and more sophisticated verbal descriptions of the music in question. No apparent effects of skill development were observed for the novice music readers alone, but all participants improved their verbal descriptions towards the last measurement. Apart from the background-related differences between groups of participants, combining verbal and eye-movement data in a cluster analysis identified three styles of silent reading. The finding demonstrated individual differences in how the freely defined silent-reading task was approached.

This dissertation is among the first presentations of a series of experiments systematically addressing the visual processing of music notation in various types of music-reading tasks and focusing especially on the eye-movement indicators of developing music-reading skill. Overall, the experiments demonstrate that the music-reading processes are affected not only by "top-down" factors, such as musical background, but also by the "bottom-up" effects of specific features of music notation, such as pitch heights, metrical division, rhythmic patterns and unexpected melodic events. From a methodological perspective, the experiments emphasize the importance of systematic stimulus design, temporal control during performance tasks, and the development of complementary methods, for easing the interpretation of the eye-movement data. To conclude, this dissertation suggests that advances in comprehending the cognitive aspects of music reading, the nature of expertise in this musical task, and the development of educational tools can be attained through the systematic application of the eye-tracking methodology also in this specific domain.

Keywords: eye-tracking, music education, music reading, sight-reading, silent reading, skill development

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

In 2002 I stepped through the gates of the University of Turku as a 19-year-old girl planning to become an elementary school teacher. Little did I know that the years at the university would change me enough to pursue a line of work I had never really thought for myself.

I wouldn't be here without support and encouragement from many people. First and foremost, I'm deeply grateful to my supervisors Professor Mirjamaija Mikkilä-Erdmann, Professor Erno Lehtinen and my former supervisor Professor Emeritus Erkki Olkinuora for their guidance during these past six years. I especially want to thank my first supervisor, Mirjamaija, for putting the idea of a doctoral dissertation into my head, encouraging me to start working on a methodology I had never heard of before, and believing in my graduation from the very beginning. As a young master student, I wouldn't have been self-assured enough to even try such things on my own initiative.

As anyone can see from the original publications included in this dissertation, there are two people who have had a tremendous role in this research project: my co-authors Professor Erkki Huovinen (University of Minnesota) and Anna-Kaisa Ylitalo (University of Jyväskylä). I don't know how to thank you two enough for offering your expertise to the project's use, or for the time and effort you have invested in all this. We have learned a lot together, and I am more than happy that all of us hope to continue our co-operation. Most of all, however, I value the fact that though we started out as complete strangers, we now have become good friends. Dissertations and publications aside, this is what I consider a profitable outcome of a research project.

There are many colleagues at the Department of Teacher Education who have helped me in the past, and also some who have already helped me to look into the future. I'd especially like to mention Pekka Salonen and Markku Silander and their input at the first stage of our music-reading project, Professor Arja Virta and Markus Nivala from our current project on history learning, and Ilona Södervik and Henna Vilppu from the on-going project "Learning of Medical Expertise". I'm lucky to have such colleagues representing so many areas of expertise. My deepest compliments also go to my other colleagues at the Department of Teacher Education and the Centre of Learning Research. It was a big thing to start working at a place where I had just been a student, but I felt welcomed.

Erkki Anto, it's been almost eleven years since we started our studies at the Department of Teacher Education – and here we still are. It's an incredible feeling to have someone walk by your side, sharing all the ups and downs, as you're trying to grow into a profession. Thank you.

I'd also like to thank my friends and relatives. You have all suffered a great deal of university jargon on my part during the past years, and offered me the all-important and welcome distractions from the university life. Mom and dad, this project on eye movements in music reading has obvious connections to both of your life-long careers. I hope it demonstrates how much I value your work.

And of course, there are the ones I love the most. At the same time I've worked on this dissertation and strived to understand what it means to be a researcher, I have also become a partner and a parent. That has really put things in perspective. Dear Toni, Odin and Milan, thank you for being there. You make me who I am.

In Turku, March 2013

Marjaana Penttinen

LIST OF ORIGINAL PUBLICATIONS

1) Penttinen, M., & Huovinen, E. (2009). The effects of melodic grouping and meter on eye movements during simple sight-reading tasks. In J. Louhivuori, T. Eerola, S. Saarikallio, T. Himberg, & P. Eerola (Eds.), *Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music*, Jyväskylä, Finland (416–424). Retrieved from https://jyx.jyu.fi/dspace/handle/123456789/20910

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

3) Penttinen, M., Huovinen, E., & Ylitalo, A. (2012). Unexpected melodic events during music reading: Exploring the eye-movement approach. In E. Cambouropoulos, C. Tsougras, P. Mavromatis, & K. Pastiadis (Eds.), *Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Conference of the European Society for the Cognitive Sciences of Music,* Thessaloniki, Greece, (792–798). Retrieved from http://icmpc-escom2012.web.auth.gr/sites/default/files/papers/792_Proc.pdf

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

5) Penttinen, M., Huovinen, E., & Ylitalo, A. (in press). Silent music reading: Amateur musicians' visual processing and descriptive skill. *Musicæ Scientiæ*.

1. INTRODUCTION

In many musical genres, the ability to read music notation is, though not a prerequisite, at least a great facilitator of performing, rehearsing, composing and arranging music (e.g., Lehmann & Kopiez, 2009; Sloboda, 1978), and the teaching of the skill is considered as an important part of music education. The methods applied in teaching music reading, however, appear to be based more on intuition and tradition than actual evidence of how people learn this particular skill (see Gudmundsdottir, 2010a). Indeed, a systematic study of the different stages of learning to read music notation as well as studies on the "model" performances of skilled musicians are needed in order to refine the current teaching methods.

The present dissertation explores the possibilities of the *eye-tracking methodology* in research on music reading. In practice, eye tracking is the measuring of the points of gaze during the inspection of a given visual stimulus, with a specially manufactured video-recording device. The method has long research traditions in text reading and to a lesser extent in scene perception and visual search (Rayner, 1998, 2009). However, even despite the early starts by Jacobsen (1928) and Weaver (1943), research on eye movements in music reading has remained underdeveloped (Madell & Hébert, 2008). Furthermore, the emphasis has been on the reading processes of skilled musicians, whereas the development of this specific skill has gained only minute attention. The series of experiments presented in this dissertation systematically apply the eye-tracking methodology in the context of music reading. More precisely, the experiments explore novices' music-reading in tasks with and without performance requirements, and the eye movements of more advanced musicians in accurate "model" performances of simple melodies.

1.1 Reading music notation

In this dissertation, *music reading* in all cases refers to the reading of *modern staff notation*, i.e. the symbolic system commonly applied in Western music tradition, where sounds are visually represented by organizing them into a (five-line) staff. Examples of one-line *musical scores* written in Western music notation are presented in Appendix 1. Unlike in auditory perception of music, music notation gives the performer specific information on many aspects of the music in question, such as rhythm, melody, harmony, and context (see Temperley, 2001; Wristen, 2005). To simplify, the rhythm, i.e. the lengths of individual notes and the piece's metrical framework, is implied by the stems and flags of individual note symbols that are then suitably fitted between vertical bar lines according to the given meter. The melody, i.e. the succession of pitch heights, is reflected in the horizontal locations of concurrent note heads. Harmony, in its part, is presented by groups of simultaneously performed notes or by chord symbols placed above the staff lines. Finally, context is given by textual information instructing to perform the piece "vividly" or "slowly", or by expressive markings. Overall, music notation may thus contain multiple types of information and, quite in detail, guide the reader to perform the music according to the composer's intentions. However, musical scores may also be simplified. The melodies in Appendix 1, for instance, do not contain any contextual information, and only the final melody presents exact information on harmony in the form of chord symbols. The complexity or the wealth of information on musical scores may thus be quite easily altered, and not all scores represent the sounds with equal precision.

1.1.1 Three types of music reading: sight-reading, rehearsed reading, and silent reading

Research literature on music reading is not entirely consistent in defining the act itself. Some incorporate musical instruments into the reading task (Gudmundsdottir, 2010a), some see music reading resulting in not only external but also, alternatively, internal musical sounds (Hodges, 1992), while others consider music reading also to contain the inspection of a musical score without performance demands (Lehmann & Kopiez, 2009). In this dissertation, music reading is used as a term covering three subcategories. First, *sight-reading* (or sight-singing) is easily described through its Italian counterpart: during a *prima vista* performance, the reader is expected to perform the piece “at first sight”. Second, the most common type of music reading, at least when performing Western art music, is *rehearsed reading*, in which a performer is reading and performing an already familiar piece of music. Third, *silent reading* includes the reading(s) of a piece of music with no simultaneous performance. The latter type of reading may occur, for example, when a musician prepares herself for a performance by scanning through a score (cp. “music reading” of Lehmann and Kopiez [2009] above). All these types of music reading are serving different purposes and are of a different nature in terms of their cognitive demands. They also present different opportunities and limitations for eye-tracking research within this domain.

Sight-reading and rehearsed reading. Reading music during a performance is, without a doubt, a complex and a challenging task. First of all, the complexity of music notation has much to do with its multidimensional nature, as the note symbols typically present information on both pitch and timing simultaneously (Gudmundsdottir, 2010b; Waters & Underwood, 1999). In addition, a performer is required to transform this complex information into correct motoric responses. Finally, to complicate the matter even further, the presence of timing information regulates the flow of motoric output. Thus, performers not only have to recognize the musical symbols and be aware of their correct motoric responses, but also execute them *at a given pace* (see, e.g., Kinsler & Carpenter, 1995.) In addition, especially on higher levels on expertise, the contextual, or expressional, aspects create even further challenges for the performer.

The difference between sight-reading and rehearsed reading lies in that during sight-reading, the performer does not benefit from prior experience with the particular musical material. During the performance of a rehearsed piece of music, on the other hand, the performer already has memory traits on the music’s structures and units, which help in preparation for the specific motoric responses. (e.g., Palmer, 1997; Sloboda, 1984.) It needs to be noted here, however, that in addition to familiarity with a particular piece of music, all types of music reading are naturally influenced by more general prior knowledge of music and, more specifically, on the type of music being performed (Lehmann & Ericsson, 1996). At the extreme case, expert musicians may become so familiar with a specific musical style and its conventions, such as the classical style defining the structure and melodic patterns of Mozart’s sonatas, that although seeing a score for the first time, the performance may not be any less perfect than later, rehearsed ones (Wolf, 1976). However, the distinction between sight-reading and rehearsed reading still holds in terms of visual processing, though the auditory outcomes of such reading tasks may, on occasion, sound similar.

The concept of sight-reading is often somewhat vaguely used (Gudmundsdottir, 2010a; Wristen, 2005). This is most likely due to the tradition according to which the act of sight-reading is considered to contain some form of preparation. For instance, Lehmann and Kopiez (2009, p. 344) define sight-reading as "non- or under-rehearsed music reading that aims at an adequate performance in terms of tempo and expression". Along with similar, generally accepted definitions, participants in many studies on sight-reading have been allowed a more or less defined time period to familiarize themselves with the piece of music by silently reading the score or even by practice with an instrument. Such a procedure is admittedly natural to musicians, who rarely have to perform with no time for preparation whatsoever. However, it has been noted that musicians' performances improve to a notable extent already during the second or third performances of a musical score, compared to the first attempt (e.g., Fine, Berry, & Rosner, 2006; Goolsby, 1994a), and that growing familiarity with the music also affects the course of the visual processing (Goolsby, 1994a; Kinsler & Carpenter, 1995). Thus, this dissertation suggests that for research purposes, it is most useful to differentiate the first encounter of the musical material (sight-reading) from its later encounters (rehearsed reading).

Silent reading. Silent reading of music notation is often seen as a subtask of a "true" music-reading task, and it has only rarely been the sole target of research (though see section 1.2.1). Silent reading differs from sight-reading and rehearsed reading in two central aspects. First, the need to transform the visual symbols into motoric actions is eliminated. Second, the demand for serial, mainly progressive reading proceeding within a pre-defined temporal framework is relaxed. Thus, the emphasis shifts from using music notation simply as a performance guide towards the conception of music notation as a source of musical information, applicable in multiple ways. In this dissertation, silent music reading is, then, simply a process of extracting information from music notation without performance requirements.

As described above, silent reading is often viewed as an essential part of the sight-reading procedure. Along with this, it has been suggested that advanced skills in sight-reading are related to how the reader prepares himself for a performance by silently reading the score before the actual performance. A more skilled sight-reader, for instance, may check the key elements (such as the key signature) of the melody, or try to search for difficult passages within the piece prior to performing, whereas a less skilled sight-reader does not apply such strategies (Drai-Zerbid, Baccino, & Bigand, 2011; McPherson, 1994, 2005). However, apart from planning a performance, the reader could also be facing a variety of other types of cognitive tasks during silent reading. Thus, silent reading should be regarded as a unique form of music reading, as it poses a variety of approaches to the visual processing of this particular type of stimuli. In cases such as the error-detection task (e.g., Gilman & Underwood, 2003), for instance, silent reading likely becomes more a task of visual search: visual cues for symbols violating the expectations created on the basis of prior knowledge of music theory and musical conventions may be searched instead of thinking how the score would be performed or how the music would sound (on this *auditory imagery*, see, e.g., Brodsky, Henik, Rubinstein, & Zorman, 2003; Brodsky, Kessler, Rubinstein, Ginsborg, & Henik, 2008).

Thus, in silent reading, the reader may not be obliged to examine all detailed information the score contains, or examine the notation at all linearly, or "hear" the music, but the visual processing is selective and guided more by the requirements of the given task. In some cases, silent reading may actually resemble more image inspection than "reading" in its strictest sense. The distinction between more or less linear image inspection and actual reading is, in practice, impossible to determine. Therefore, the concept of silent "reading" is considered here broadly, and it covers all types of inspections of music notation without performance requirements. Considering the fact that domain knowledge in the context of music is quite conveniently measured, silent reading offers an interesting area for eye-tracking research in non-textual domains.

In the experiments presented in this dissertation, all three types of music reading are examined, and suitable measures targeting the key aspects of each of the music-reading tasks are proposed. In Experiment I, participants performed a true *prima vista* task, as the short melodies were composed for the experiment alone, and no additional time for practicing or inspecting the melodies beforehand was offered. In Experiment II, the stimuli were the melody of a well-known children's song and its simple variations. The participants were allowed to practice the original melody; the task thus resembled rehearsed reading, while the variations offered a way to examine the adjustment of visual processing in the face of unexpected melodic events. In Experiment III, a task of silent reading was administered, and participants were asked afterwards to verbally describe what they had seen in the musical image.

1.1.2 Development of music-reading skill

Most of the knowledge attained so far on music reading is based on studies conducted with expert music readers, whereas research on the development in this domain is rarer (see Gudmundsdottir, 2010b). To begin with, expertise in music reading shares many features with expertise in other domains. Above all, experts in music are also able to effectively code and store musical information by "chunking", that is, by organizing multiple pieces of information as a single piece of information (Chase & Simon, 1972; de Groot, 1965; Lehmann & Gruber, 2006; Sloboda, 1978). In the case of tasks with high cognitive load, such as sight-reading, the ability to chunk information is presumably a great benefit. In practice, during music reading, an experienced performer's larger musical "dictionary" eases the recognition of musical patterns (Waters, Underwood, & Findlay, 1997). It has indeed been noted that more proficient sight-readers are able to more efficiently recognize familiar melodic and rhythmic patterns from music notation than less proficient ones (e.g., Burman & Booth, 2009; Fine et al., 2006; Waters & Underwood, 1998; Waters et al., 1997; Wolf, 1976).

A curious feature of expertise in music reading is that the skills of sight-reading and rehearsed reading, despite their general positive correlation during the earlier stages of skill development (McPherson, 1994; 2005), do not necessarily go hand in hand in the case of high-level experts (Wolf, 1976). According to Lehmann and Ericsson (1996), task-specific training, i.e. experiences with sight-reading situations and increase in musical repertoire, generally promote expert performers' sight-reading ability. However, it has also been suggested that sight-reading ability is not only determined by practice dependent skills but also by innate abilities such as psychomotor speed (Kopiez & Lee, 2006) or working memory capacity (Meinz & Hambrick, 2010). For non-expert musicians, possible background factors are reading comprehension, rhythmic audiation, visual field articulation, spatial orientation (Gromko, 2004), or reading and comprehending rhythms (McPherson, 1994). The general role of notational audiation in sight-reading, and whether it is the cause or the result for enhancing various types of musical skills, is also still subject to debate (see Hubbard, 2010). All in all, the multifaceted issue of expertise in music reading would offer means for studying a variety of psychological topics within this specific domain (Lehmann & Gruber, 2006).

Although little is known on skill development in music reading from the perspective of visual processing, some studies on musical performance have addressed the issue. First, in their comprehensive study on the effects of skill and practice for novice, intermediate and expert performers, Drake and Palmer (2000) reported that performance tempo and accuracy as well as the ability to maintain temporal continuity improved with skill level and practice, while the range of planning increased only with skill level. In addition, some findings suggest that performers are sensitive to the presence of metrical boundaries in music notation, as correction errors tend to occur more likely at these boundaries (Drake & Palmer, 2000; Palmer & Drake, 1997; see also Sloboda, 1984). The visual processing of music notation can be expected to develop parallel to such changes in performances. For instance, one may assume that metrical bar lines play a role in organizing visual processing, or that some effects of skill are demonstrated in eye-movement patterns related to the planning of motoric responses (as indeed seems to be the case: see section 1.2.1 below).

Second, studies have demonstrated that in cases where an errorless musical performance is not obtainable, some elements of the information are preferred over others. In general, the focus tends to be placed on pitch information at the expense of rhythmic accuracy. This holds for both beginners and experienced musicians with less advanced skills in sight-reading. (Gudmundsdottir, 2010b; Drake & Palmer, 2000; Henry, 2011; McPherson, 1994; Pike & Carter, 2010.) These findings suggest that information on pitch and rhythm (typically presented by any one note symbol) is processed serially at least in some stages of developing music-reading skills. However, it is still an open question whether the processing of pitch and rhythm share some of the cognitive mechanisms or not (e.g., Waters & Underwood, 1999). Henry (2011) suggests these findings to be a consequence of training within the Western music tradition, which emphasizes the importance of producing “correct” sounds. This appears a likely explanation. However, in terms of examining less experienced performers’ visual processing during music reading, this matter poses a problem for data analyses: disruptions in the temporal progression alter the temporal framework so that some sections of the music are simply getting more attention than others, and the analyzed performances thus become incommensurable with regard to the use of time. This issue would be overcome by focusing on performances of similar duration.

The performance approach to skill development in music reading interprets development in this area to be first and foremost manifested as the lessening of performance errors. Examining the causes of errors gives clues to what may hinder the reading, but the approach is limited, as all of the processes involved in *fluent* music reading remain hidden. The main methodology applied in the present dissertation, eye tracking, offers one alternative for gaining exact data on the course of visual processing during various types of music-reading tasks. Furthermore, the performance approach necessarily connects reading ability to performance ability, though it is not certain that these abilities develop completely in parallel. Perhaps in some cases, music notation is comprehended better than it can be executed. It is therefore proposed in this dissertation that meaningful musical information may also be extracted from a score in silent-reading situations.

1.2 Eye movements in music reading

The established view concerning visual information processing is that information is gathered during short moments when the eyes are relatively still, i.e. *fixations*. An eye movement between two fixations is called a *saccade*. Saccades are so quick (30–50 ms) that visual information is suppressed during these rapid shifts (Rayner, 1998, 2009). In principle, then, music reading, similarly to all eye movements, consists of a series of fixations and saccades, each saccade moving the narrow spectrum of accurate vision to a new target location. The basis for investigating eye movements lies in the general observation that viewers' fixation durations and saccade lengths under different visual tasks (such as text reading, scene perception, or visual search tasks) do not correlate. Thus, visual processing is not simply the decoding of information in a "bottom-up" fashion, but also "top-down" processes, i.e. cognitive mechanisms, must influence the process of information search (Neisser, 1976; Rayner, 2009). Though the eye-tracking methodology does not give direct access to such cognitive mechanisms, it is a useful method for collecting on-line data on information processing. The following sections review eye-tracking research on music reading since the 1990s, and discuss some limitations of the previously applied research designs.

1.2.1 Skill-related differences in eye movements during music reading

Figure 1 presents a skilled adult amateur musician's fixations during the performance of a familiar tune. During a single fixation, readers process information available within their *perceptual span*, which in music reading extends 3–5 beats (Gilman & Underwood, 2003; Truitt, Clifton, Pollatsek, & Rayner, 1997) or 4–5 notes (Burman & Booth, 2009) to the right from the point of the fixation. There have so far been no reports that the size of the perceptual span would reduce or increase as a function of music-reading skill, though such effects have been reported in the context of text reading (see Rayner, 2009). It has been suggested, however, that perceptual span may increase due to rehearsal, at least in an error-detection task (Burman & Booth, 2009). The reading of music typically proceeds with a chain of progressive saccades moving along the musical score. Occasionally, though, this progression is disrupted, and the forward inspections are followed by regressive saccades (Goolsby, 1994a, 1994b; see Figure 1).

Figure 1. Example of an adult amateur musician's succession of fixations during an errorless piano performance of bars 2 to 7 of "Mary Had a Little Lamb" in the tempo of 60 bpm (Experiment II of the present dissertation). Each arrowhead refers to the exact location of a fixation. Note the regressive eye movements (fixations number 7 and 14). Image by Anna-Kaisa Ylitalo (2012).

Some studies have aimed to clarify which features of music notation are, in fact, the exact targets of fixations. In Figure 1, for instance, the skilled performer does not fixate on all notes, and the fixations do not solely land on the note heads, as one might expect. Occasionally, fixations seem to target spaces, note stems or note tails, and overall, they appear to follow the melodic contour, moving upward in the vertical axes as the melody goes "up", and downward after bar number four, where the highest notes of the melody are located. Similar observations on fixation locations have been previously made by Gilman and Underwood (2003), Goolsby (1994b), and Truitt and colleagues (1997). However, the reported amount of time spent on fixating note heads, for instance, varied in these studies from 33 % (Truitt et al., 1997) to 58 % (Gilman & Underwood, 2003). Thus, it is not yet apparent how significant a role the exact fixation location has during music reading. What was common to the studies of Gilman and Underwood (2003) and Truitt and colleagues (1997), however, was that bar lines were scarcely fixated (1 % and 7 %, respectively). This suggests that the fixations are not randomly targeting *any* one symbol on the score. Indeed, the several detailed examples of two performers' eye movements reported by Goolsby (1994b) show that in general, fixations appear to target, if not precisely on the note symbols, at least *around* them.

Thus, the studies above do not suggest that particular parts of a note symbol would be, in all cases, the preferred targets of attention or information extraction. One explanation may lie in performance tempi: in faster tempi, saccades tend to lengthen, as the gaze moves more quickly across the stimuli (Kinsler & Carpenter, 1995). Perhaps variability in performance tempi has slightly confounded the findings mentioned above, if when in haste, performers' fixations have not landed exactly on the intended items. In addition, at present, no studies have examined the *vertical* size of the perceptual span in music reading. In tasks of visual search during text reading, information even from below the currently fixated line may be obtained (see Rayner, 2009). This suggests that the area of accurate vision may also in music reading cover an area wide enough to contain all of a note symbol (depending on the size of the score), despite the exact fixation location. Overall, then, though the note symbols consist of separable parts with different music-theoretical meanings, they are perhaps best to be examined as "wholes", as far as visual processing is concerned. The different roles of the various aspects of the symbols, such as pitch and rhythm, may then be explored by simply controlling the stimuli. The present dissertation is based on the assumption that within the limitations of performance tempo, the gaze most often targets *as close as possible* to an intended visual item. Therefore, the experiments adapt the style of assigning fixations to their nearest note symbols (cp. the "lenient" scoring criteria of Truitt et al., 1997). In Experiment I, data analyses will be based on note-specific fixation times; in Experiment II (see Figure 1), on quarter-beat areas.

In music reading, the average fixation duration has been reported to lie between 200 and 400 ms (Madell & Hébert, 2008; Rayner, 1998), though this is subject to variance between individuals as well as music-reading tasks. It is also subject to great standard deviations; for instance, reported fixation durations during a performance task have extended even beyond 1500 ms (Goolsby 1994b). In general, fixation durations are considered to reflect the time and effort needed to process the fixated information; during text reading, for instance, longer or unfamiliar words tend to be read with longer fixation durations than is the case for short and familiar words (Just & Carpenter, 1980). Accordingly, skill improvement is related to the gradual use of shorter fixation durations (e.g., Ashby, Rayner, & Clifton, 2006; Häikiö, Bertram, Hyönä, & Niemi, 2009; McConkie et al., 1991). Similarly, musicians with higher skill levels in sight-reading have been generally found to operate with shorter fixation durations than less-skilled sight-readers do, both during performance tasks (Goolsby, 1994a; Truitt et al., 1997; though note the lack of such finding in Gilman & Underwood, 2003) and silent reading (Waters & Underwood, 1998; Waters, Underwood, & Findlay, 1997).

In addition, increasing sight-reading skill apparently facilitates processing visually distinct and/or simple groups of notes by single fixations (Goolsby, 1994b; Kinsler & Carpenter, 1995; Polanka, 1995; Waters et al., 1997). Figure 1 suggests also that this reader is, indeed, fixating only once on some pairs of eighth-notes. It may be that groups of notes are actually processed as intervals, i.e., identifying a relationship between two notes rather than two separate, individual symbols (Lehmann & Gruber, 2006). Note, however, that the fixation location itself does not directly reveal such cognitive strategies. Be it as it may, this "visual" chunking of information seems closely related to cognitive chunking of information characteristic of expert musicians (see section 1.1.2 above). Correspondingly, less advanced music readers may fixate more frequently on individual note symbols than more advanced readers do (Goolsby, 1994b), although no significant differences have been reported so far in the average saccade lengths of more and less advanced music readers (see Gilman & Underwood, 2003; Goolsby, 1994a; Truitt et al., 1997).

When performing notated music, the gaze moves along the musical score slightly ahead of the current point of execution. The concept of *eye-hand span* – variously measured in time, number of beats, spatial distance, or number of notes – is used to describe this difference between the executed note and the currently fixated one (see Madell & Hébert, 2008). Truitt and colleagues (1997) suggested that a performer's gaze tends to remain even surprisingly close to the current point of execution, in their case, slightly over one beat. Even between note onsets, the gaze may be only slightly further ahead of the hands on the musical score (Truitt et al., 1997). Nevertheless, more proficient sight-readers appear to operate with larger eye-hand spans than less proficient ones do, given that the span is measured in the number of notes, beats, or spatial distance (Furneaux & Land, 1999; Gilman & Underwood, 2003; Truitt et al., 1997; see also Goolsby, 1994b). When calculating the *time* lag between the gaze and performance, however, the average length of the span may typically be around one second (Furneaux & Land, 1999; Wurtz, Müeri, & Wiesendanger, 2009), though slower performance tempo may slightly increase the length of the span (Furneaux & Land, 1999).

Finally, the eye-movement effects of structural features of music notation have scarcely been addressed. So far, we only know that increasing musical complexity generally seems to decrease the size of the eye-hand span (Gilman & Underwood, 2003; Truitt et al., 1997; Wurtz et al., 2009). The stimulus complexity has also affected fixation durations and saccade lengths for more experienced readers during a pattern-matching task (Waters & Underwood, 1998; Waters et al., 1997). Ahken, Comeau, Hébert, and Balasubramaniam (2012) noted that incongruent musical endings of short melodies were processed with longer fixation durations and more fixations than congruent endings, the effects being similar to reading of more or less plausible endings to written sentences. However, the unorthodox presentation of a key signature by using accidentals in front of individual notes removed the effects of incongruence (Ahken et al., 2012). Overall, the findings suggest that also the fine-grained properties of music notation should be taken more into account when examining eye movements in music reading (see Madell & Hébert, 2008).

1.2.2 Methodological considerations emerging from prior studies

The prior eye-tracking studies on music reading have included certain tasks of silent reading, such as error detection (Burman & Booth, 2009; Gilman & Underwood, 2003), "reading as if sight-reading" (Polanka, 1995) and pattern-matching (Waters & Underwood, 1998; Waters et al., 1997). Others have included performance tasks such as the tapping of notated rhythms (e.g., Kinsler & Carpenter, 1995) or reading actual musical works while singing, playing an instrument or even conducting an ensemble (e.g., Adachi, Takiuchi, & Shoda, 2012; Ahken et al., 2012; Bigand et al., 2010; Drai-Zerbib et al., 2011; Furneaux & Land, 1999; Gilman & Underwood, 2003; Goolsby, 1994a, 1994b; Truitt et al., 1997; Wurtz et al., 2009). The studies are quite few in number, at least in comparison to research in other domains, and pose some methodological issues that future studies should address.

Summarizing previous findings effectively is difficult due to the variability of the applied musical stimuli. The selection of sight-reading tasks in the studies has tended to emphasize ecological validity by making use of real, complex pieces of music (e.g., Bigand et al., 2010; Burman & Booth, 2009; Drai-Zerbid et al., 2011; Furneaux & Land, 1999; Wurtz et al., 2009), although sometimes in a simplified form (Gilman & Underwood, 2003; Truitt et al., 1997). More or less systematic stimulus designs have appeared mainly in studies on silent reading (Polanka, 1995; Waters & Underwood, 1998; Waters et al., 1997; see also Adachi et al., 2012). Finally, in some studies, the musical material has even differed for readers with more or less musical experience (Burman & Booth, 2009; Furneaux & Land, 1999). The careful controlling of stimuli has long been a standard in research on text reading, and there is no reason not to strive for a similar level of precision in music-reading research as well, especially when considering the complexity of music notation. The challenge lies in maintaining the meaningfulness of the tasks (in a musical sense), so that performers act similarly to typical music-reading situations.

In addition, music studies have been typically conducted with relatively few participants (four in Kinsler and Carpenter [1995]; seven or eight in Furneaux and Land [1999], Truitt et al. [1997], and Wurtz et al. [2009]; 18 in Adachi et al. [2012], Ahken et al. [2012] and Polanka [1995]; 22 in Waters and Underwood [1998]; 24 in Waters et al. [1997]; and 25 in Drai-Zerbid et al. [2011]). Quite naturally, the greatest reason for this has been the effortful eye-tracking equipment applied in the 1990s, which also lacked the accuracy of modern recording devices. With modern equipment, it is remarkably easier to both collect and handle larger data samples. Another reason for the limited amount of participants may be that studies have often focused on high-level experts or very skilled non-professionals, and such participants are often hard to find. Furthermore, the participants of prior studies have also been typically divided into 2–3 skill-based groups. Music skill has, however, been defined differently across studies: the group divisions have been based either on participants' music-reading ability (Gilman & Underwood, 2003; Goolsby, 1994a; Polanka, 1995), performance ability (Drai-Zerbid et al., 2012; Truitt et al., 1997) or more general musical experience (Burman & Booth, 2009; Furneaux & Land, 1999). Therefore, it is not guaranteed that the skill-based groups of one study entirely match the groups of others. Studies applying *complete* novice music readers have apparently been reported only by Adachi and colleagues (2012) and Waters and colleagues (1997; though see the control procedure in Burman & Booth [2009]). It is suggested here that by using simplified stimuli, it is not necessary to only have the best of performers as study participants, but one can also select performers that are experts *in* *relation* to the difficulty of the stimuli. The observed consistency in some of the discoveries in the prior studies, then, should be corroborated with larger amounts of data, while future studies also should carefully consider the selections of study participants.

To some surprise, a key aspect of music reading, namely *temporal control*, has seldom been an issue in the prior studies. Instead, performers have been allowed to choose tempi of their own preference (with the notable exception in the study by Kinsler and Carpenter [1995] as well as partial temporal control in Furneaux and Land [1999]). As less-skilled music readers have reportedly chosen slower tempi than more-skilled readers (Gilman & Underwood 2003; Truitt et al. 1997), the observed skill-related effects in eye movements can be, in part, a consequence of some participants having more time to perform a piece than others (for a similar remark, see Lehmann & Kopiez, 2009). Indeed, Wurtz and colleagues (2009) reported a positive correlation between the time violinists needed for playing a piece and the time spent on visual anticipation while reading the score. Another factor related to temporal control is the fact that during a performance, even skilled readers do not need the same amount of time and effort to process each musical symbol and transform it into motoric actions (see Kinsler & Carpenter, 1995), but some items should be more difficult and others easier to process and execute. This is especially relevant in performances controlled by a fixed tempo, in which the confronting of more challenging notational information should increase the cognitive strain for the performer, who then needs to adjust the reading of the musical score in such a way that a continuous (and preferably correct) performance is possible to obtain. On the eye-movement level, coping with this additional cognitive strain might be reflected in multiple ways, but this could only be fully examined with a pool of temporally commensurable performances. So far, little information is available on this issue. All in all, to increase systematics in this line of research, it is of utmost importance to perform studies in controlled tempi and thus examine performances alike with regard to the use of time.

To develop this area of research even further, a more systematic approach to analyzing the eye-movement data is needed. Instead of the traditional approach of defining fixations simply as progressive or regressive, the present dissertation applies measures used in eye-movement studies on text reading (see Hyönä, Lorch, & Rinck, 2003). Apart from the studies presented in this dissertation, these measures seem to have been applied only once in the context of music reading, by Drai-Zerbid and colleagues (2011). In addition, prior eye-movement analyses in this domain have typically been based on absolute fixation durations. This is somewhat problematic, when performance durations have varied. In this dissertation, where also the visual processing of novice performers is examined, some of the eye-movement analyses explore the allocation of fixation time during music reading from a slightly different perspective: the focus is on the distribution of fixation time across different musical symbols. This approach is based on the assumption that some symbols might collect relatively more fixation time than others despite differences in performance durations or performers' music-reading ability.

The planning of motor responses during music reading has been examined almost solely through the size of the eye-hand span (but see Truitt et al., 1997). The eye-hand span is, of course, one of the main tools for adjusting the reading of a musical score so that a temporally correct and accurate performance is possible to obtain. However, the eye-hand span is only a measure of distance defined for individual points in time, and does not tell us how widely the score is being scanned during a given time interval. For instance, given that beginning performers operate with longer fixation durations, such a performer might not have time for more than one or two fixations during the performance of one metrical beat in a temporally controlled condition. A more accomplished music reader, on the other hand, operating with shorter fixation durations, might be able to scan a broader area of the score during the same time. Indeed, it has already been suggested that skilled music readers make use of the available time – especially when performing notes with longer durations – by inspecting upcoming musical material (Goolsby 1994a, 1994b; see also Truitt et al. 1997). In the present dissertation, a measure entitled *gaze activity* is introduced to reflect the amount of inspected musical information between the onsets of two quarter-beats. In addition, contrary to prior studies, the eye-hand span will be explored in temporally controlled conditions, with a precise synchronization of performance and eye-movement data, and with reasonably sized participant groups.

As Lehmann and Kopiez (2009) rightly suggest, for the purposes of future research, more attention should be placed on the development of standardized study conditions. Apart from the issues described above, this dissertation proposes also the differentiation between reading with or without performance requirements as one such aspect of development. The present dissertation also aims to highlight the role of music notation as a source of musical information beyond its well-known role as a performance aid. A problem yet to be solved, however, is the difficulty of measuring reading efficiency during silent reading. This dissertation offers one alternative by analyzing readers' retrospective verbal descriptions of the music after silent reading. Using complementary verbal reports in conjunction with eye-tracking data is not uncommon in eye-movement research in general. Recent eye-tracking studies have applied both think-aloud procedures (e.g., Kaakinen & Hyönä, 2005; Kendeou & van den Broek, 2005; 2007) and the cued retrospective reporting technique, in which participants are interviewed while being presented with a video recording of their own fixations during stimulus inspection (e.g., Hansen, 1991; Hyrskykari et al. 2008; Jarodzka, Scheiter, Gerjets, & van Gog, 2010; Penttinen, Anto, & Mikkilä-Erdmann, 2012; van Gog, Paas, van Mërrienboer, & Witte, 2005). In the present dissertation, however, the focus is on readers' descriptions of a musical stimulus after silently reading a score, and not on their thought processes. For this reason, in one of the experiments of the present dissertation, retrospective reports without video cuing were conducted. In the context of eye-movement research, such non-cued retrospective descriptions were notably applied in de Groot and Gobet's (1996) study on chess players' cognitive processes, based on data collected in the 1960s. In their study, brief visual exposition of chess positions was followed by letting the participants reproduce the positions on a chessboard and retrospectively report on their thought processes. De Groot and Gobet (p. 200) also paid attention to the relative richness of the semantic categories inherent in the verbal protocols, classifying the reported "percepts" with a view to whether or not they involved dynamic relations between the chess pieces or references to abstract move possibilities. In a similar vein, the present dissertation introduces two kinds of measures for assessing the quality of silent readers' retrospective descriptions of a musical score – an *accuracy* measure as well as a measure of semantic richness, to be called *integration*.

2. MAIN AIMS

The general objective of this dissertation is to explore eye-movement indicators of skill and skill development in music reading and to develop and test suitable methods that can be used by future studies on the topic. The dissertation thus aims to promote two major issues.

1) The first aim is to examine *the eye-movement indicators of skill and skill development in various music-reading tasks*. First, in Experiments I and III, changes in beginning adult music readers' sight-reading and silent reading across a nine-month long music-training period were examined. Amateur musicians constituted a point of reference for assessment of the novices' development. Second, in Experiment II, the visual processing of music education students was compared to that of expert performers during temporally controlled rehearsed reading, focusing here on the eye-movement effects of unexpected melodic alterations placed within a rehearsed melody. In Experiment I, the data analysis exploited the eye-movement measures traditionally applied in text-reading research; in Experiment II, two types of eye-hand span measures were examined, introducing a measure entitled gaze activity; and in Experiment III, a two-way categorization for verbal descriptions was developed for the purpose of qualitatively evaluating the information extracted from music notation in a non-performance task, and verbal data was investigated in combination with selected eye-movement measures.

2) The second aim is to *design and test research settings that can be used by future studies on eye movements during music reading.* Considering the tasks of sight-reading and rehearsed reading, special emphasis was placed on systematic stimulus design and temporal control of the performances. In Experiment I, the eye-movement effects of metrical structure and large melodic intervals were examined through the sight-reading performances of systematically constructed, simple melodies, and analyzed together with performance data. In Experiment II, visually non-salient melodic alterations of a familiar melody were created, and the research setting was further developed by a precise synchronization of the eye-movement and performance data. Finally, Experiment III examined the effects of musical experience on silent reading, and a free-viewing task of music notation combined with a post-viewing verbal protocol was administered.

3. METHODS

This dissertation consists of three experiments. The first addresses novices' eye movements during sight-reading, the second, experienced performers' visual processing during rehearsed reading, and the third, silent reading of music in a free-viewing task. In total, five scientific studies were written based on the three experiments (see Table 1).

3.1 Experiment I: Eye movements and early skill development in sight-reading

In Experiment I, the main goal was to examine eye-movement behaviour of adults during their first steps of learning to read music notation. The experiment spanned a nine-month long music training period, during which the participants sight-read and performed a series of simple, systematically created melodies in three measurement sessions.

Participants. The 49 participants were future elementary school teachers (BA students) from a Finnish university. The participants were selected from a pool of 2nd-year students taking part in a nine-month long compulsory music training period, including courses on music didactics, school instruments, and altogether 20 hours of piano lessons in small groups. In order to minimize the effects of different teaching styles and methods, all participants attended piano lessons given by the same music teacher. The musical background of the participants was assessed on the basis of written questions administered before the first measurement. Participants who reported to be unable to read music notation or who did not have any experience in playing instruments were defined as novices in music reading. Participants classified as "less experienced (LE) amateurs" reported being able to read music notation and had taken instrument lessons for less than 6 years. "More experienced (ME) amateurs", on the other hand, reported understanding music notation and had taken instrument lessons for 6 years or more.

Stimulus materials. Simple five-bar melodies (see Appendix 1) were created for the experiment with the following criteria. First, considering the novice participants of the study, the melodies had to be kept as simple as possible. Thus, all melodies were written in C major and played only with five white keys, eliminating the need to move one's hand while performing. In addition, the fingering for the first note (C4 or G4) was marked with a number (1 or 5) to ease the beginning of the performance. The melodies only contained quarter notes apart from the final whole note, eliminating the need to identify different note values, and the performance tempo was a relaxed 60 bpm (= beats per minute, i.e. in this case, one key press per second). Finally, the melodies were primarily stepwise, to ease the recognition of the notes. Second, the 12 stimulus melodies were created also with strict music-structural criteria in mind. Each melody ended on the pitch E4 and had a melodic range from C4 to G4. Within each of the primarily stepwise melodic movements, there were two larger intervals – a perfect fourth and a perfect fifth – at the temporal distance of four quarter-beats from one another. Two types of melodies were created; the intervals larger than a semitone occurred either at bar lines, or were shifted to occur two beats later, to the middle of bars 3 and 4. Each melody x had an inversion y (see Appendix 1). The melodies were previously unknown to the participants, providing an authentic sight-reading task in each measurement.

Apparatus. Eye movements during playing were recorded using a Tobii 1750 Eye Tracker manufactured by Tobii Technology AB (Stockholm, Sweden). The infrared cameras tracing the position of the participants' pupils were integrated into the body of the same computer monitor from which the stimuli were presented. Both eyes were tracked with a frame rate of 50 Hz, and the accuracy of the recording system was 0.5 degrees. The screen resolution was 1024 × 768 pixels. The eye tracker allowed slight head movements, and no chin or head supports were used. The participants performed on a Yamaha electric piano, and performances were recorded using sequencer software (Power Tracks Pro Audio).

Procedure. Each participant took part in three different measurement sessions, the first of which took place in the beginning of their music training. The second session was conducted after the second of the 8-week periods when the students had participated in circa 16 hours of group piano lessons and completed a course on elementary music theory. The third session took place at the end of the academic year, that is, after about 4 more hours of group piano lessons, the skills test, and the group music lessons. The piano course focused on improving accompanying skills, while the material differed according to the skill levels of the teaching groups; however, at the end of the course, all took part in a skills test with the same minimum requirements such as the ability to produce right-hand melodies, practiced beforehand from music notation.

Participants were tested individually in a laboratory, in the presence of an experimenter. At the start of the session, they were introduced to the equipment – an electric piano and a computer screen behind it – and then asked to adjust the piano seat to a comfortable height. Next, the eye tracker was calibrated, after which the participants were instructed to sight read four short melodies on the piano. Each notated melody (see Appendix 1) was separately displayed on the computer screen concurrently with the first metronome beat. On the screen, the width of the staff for each notated melody was 31.0 cm (12.2 in.), and the height 1.4 cm (0.55 in.). The computer screen was located behind the electric piano, the distance from the participant being about 60 cm (ca. 24 in.), corresponding to how sheet music would be placed in normal playing situations. The participants were instructed to wait for the four initial metronome beats, after which they were to start playing.

Data analysis. Errors made in the piano performances were analyzed by examining the sequencer files. The correctness of the performed pitch sequences was estimated by identifying the larger intervallic skip (which was relatively easy even in the most defective performances) and noting any pitch errors for the notes participating in these large intervals: for each recorded performance, then, the four notes participating in the two skips (first and second notes of both intervals) were examined. The performances were also analyzed in terms of their temporal characteristics. For this purpose, each performance was roughly quantized by letting the sequencer program round the time values to the nearest sixteenth note (given that bpm = 60). On this basis, simple analyses of the correctness of tempo as well as rhythmic accuracy at the larger melodic intervals were conducted. For instance, the correctness of overall tempo was estimated in terms of the inter-onset-interval (IOI) of the first notes in bars 2 and 5, expressed in the number of sixteenth notes. In order to bypass any possible effects due to beginning and ending of the melody, this analysis – as well as the eye-movement analyses – was restricted to bars 2, 3, and 4.

The eye-movement analysis applied ClearView 2.7.1 analysis software by Tobii Technology AB. Gaze remaining within a 30-pixel radius for 60 ms or more constituted a fixation. So-called Areas of Interest (AOI) with a constant width of 43 pixels were created around each note in each stimulus melody. A fixation was automatically defined into one AOI, and all statistical analyses concerning the eye movements were based on the fixation data provided this way. A measure of *overall fixation time* for each performance was defined as the total duration of all fixations beginning from the first fixation in bar 2, and ending before the first fixation to bar 5. All of the measures used in the eye-movement analysis were based on percentages of the overall fixation time. Thus, the total fixation time for a given note, as well as its components – first-pass fixation time and look-back time – were treated as percentages of the overall fixation time for measures 2, 3, and 4. The applied eye-movement measures are described in Table 1.

Table 1. Fixation, saccade and eye-hand span measures as applied in Experiments I, II and III.

	Measure	**Description**	**Experiment(s)**	**Studies**
Fixation measures	Average fixation duration	(in milliseconds)	Experiments II and III	Studies 3, 4, and 5
	Total fixation time	Summed duration of all fixations landing on a selected Area of Interest (AOI) (% of overall fixation time)	Experiments I and II	Studies 1, 2, and 3
	First-pass fixation time	Summed duration of all fixations landing on a selected AOI when it is *first* entered and until the first exit from the area (% of overall fixation time)	Experiment I	Studies 1 and 2
	Second-pass fixation time = Look-back time	Summed duration of all fixations landing on a selected AOI during its *later* inspections (% of overall fixation time)	Experiment I	Studies 1 and 2
Saccade measures	Saccade length	The pixel distance of two successive fixations' (x, y) coordinates	Experiment III	Study 5
	Incoming saccades	Frequency of saccades entering a selected AOI	Experiment II	Study 3
	Linear scanning	% of saccades proceeding between two fixations placed on the same horizontal strip (= stave) so that the latter fixation has a larger x-coordinate value	Experiment III	Study 5
Eye-hand span measures	Eye-hand span	The distance in quarter-beat areas between the currently executed quarter-beat and the fixated quarter-beat	Experiment II	Study 4
	Gaze activity	The number of fixated quarter-beat areas during the inter-onset interval (IOI) of one quarter-beat	Experiment II	Study 5

3.2 Experiment II: Effects of melodic alterations during temporally controlled rehearsed reading

Experiment II contained a pilot study before the experiment proper. Eye movement behaviour and performances of skilled musicians were recorded, while the participants read and performed the original version and variations of a familiar tune. The main aim was to examine the eye-movement effects of melodic alterations during errorless, temporally commensurable performances.

Participants. Five experienced amateur musicians from previously described Experiment I took part in the pilot study. In the experiment proper, the 40 participants represented two groups with different musical backgrounds and were (a) music education students (future elementary school teachers majoring in education) at the Department of Teacher Education of a Finnish university and (b) students of music performance at a Finnish Arts Academy or conservatory (including one participant who had completed her studies at the conservatory).

Stimulus materials. In the pilot study, three variations of a well-known children's song, "Mary Had a Little Lamb" in the key of C major were used as stimulus materials. In the main experiment, the participants performed the original version and two of the variations created for the pilot study (see Appendix 1). To create a variation, we selected one bar from the original melody in which the first note of the bar proceeded stepwise up from the previous note. Subsequently, all of the notes in this "target bar" were moved one step down so that *the first note of the target bar repeated the last note of the previous bar*. Hence, the target bar neither presented a salient visual complexity in relation to its immediate context nor did it greatly deviate in its general visual appearance from the original version of the melody.

Apparatus. A Tobii 1750 Eye Tracker by Tobii Technology AB (Stockholm, Sweden) was used to record the participants' eye movements in the pilot study (see Experiment I). In the experiment proper, eye movements were recorded using a Tobii TX300 Eye Tracker by Tobii Technology AB (Stockholm, Sweden). Both eyes were tracked with a sampling rate of 300 Hz. The screen resolution was 1920 x 1080 pixels. In both studies, a Yamaha electric piano and sequencer software (Power Tracks Pro Audio) were used to record the performances.

Procedure. In the pilot study, each participant took part in three different measurement sessions in the course of their nine-month long music course. The longitudinal aspect of the data set was not considered in the pilot study, as the participants' extensive musical training made their performances equally accurate in all three measurements and the main purpose was to develop a research setting for the main experiment. The participants were tested individually in a laboratory. At the start of the session, they were introduced to the equipment and asked to adjust the piano seat to a comfortable height. The computer screen was located behind the electric piano, distance from the participant being about 60 cm (about 24 in.). After calibrating the eye-tracker, the participants were accustomed to the setting by letting them play short diatonic melodies in C major composed by a member of the research group. Next, the experimenter performed the original version of "Mary Had a Little Lamb" in time with a metronome set again at 60 bpm, and asked whether the participant recognized the melody (all of the five participants were familiar with it). The participant was then instructed to play according to the music that would appear on the computer screen and wait for four metronome beats – after the appearance of the melody – before starting the performance. In the first measurement each participant performed Variation 1, in the second Variation 2, and in the third Variation 3.

In the main experiment, after a participant had answered written questions on his/her musical background, the experiment began with a *familiarization phase* in which the participant was first presented with the original version of the melody "Mary Had a Little Lamb" on the computer screen and asked to perform it in time with a metronome set at 60 bpm, using the right hand only. This was followed by a *practice phase* with the purpose of introducing the participant to the research protocol. During this phase, the participant was presented with a series of written instructions followed by short, simple melodies (in C major) on the computer screen while his/her eye movements and performances were recorded. Before each melody, the participant was instructed to look at a cross, marking the location of the first note two seconds in advance of the appearance of the melody. After a melody appeared on the screen, the participant was instructed to wait for two more beats (seconds) before initiating the performance. The experimenter changed all images by pressing the space bar, synchronizing her actions with a metronome set at 60 bpm to maintain a feeling of the rhythm throughout the experiment. No chin rest was used, in order to create as typical a performance situation as possible. Finally, in the actual *test phase*, the participant was first informed that s/he would next perform four versions of "Mary Had a Little Lamb" in the tempo of the familiarization phase, and that some of the melodies would contain slight alterations to the original melody. The participant was asked to perform the melodies (two original versions of the piece, A1 and A2, and variations B and C: see Appendix 1) as seen on the screen. The research protocol followed the procedure of the previous practice phase (see above). The order of the melodies was changed between every successive participant from A1–B–A2–C (Condition 1) to A1–C–A2–B (Condition 2). The assignment of the participants to one of the two conditions was randomized by allowing the participants themselves to choose an appropriate time for their session. After the test phase, the participant was presented with a video recording of his/her eye movements, and the sight-reading behaviour was informally discussed between the experimenter and the participant. For most participants, the whole session lasted between 25 and 35 minutes.

Data analysis. The five participants of the pilot study were chosen because of their flawless and temporally accurate performances; thus, no analyses of performance errors were necessary. ClearView 2.7.1 analysis software was applied in the data analysis of the pilot study. A fixation was defined as an event during which the gaze dwelled within a 40-pixel radius for 60 ms or more. Only bars 2–7 were included in the analyses, and relative total fixation times and the frequencies of incoming saccades were calculated for each of the bars, separately for each melody and participant.

In the main experiment, melodies A2, B and C were included in the analyses, as the melody A1 was merely meant to orient each performer to the task. First, performance errors were analyzed using MIDI information, differentiating between three error types: (i) note substitution, (ii) note addition, and (iii) late note (defined as at least 200 ms late with respect to the metronome). No other significant kinds of errors were observed (e.g., note omissions without substitution). The total of 912 performed bars in the data set (melodies A2, B and C) included 13 bars (1.4 %) with one or more of the above-mentioned errors. Erroneous performances were excluded from the analyses.

Second, a fixation was defined according to the default setting of Tobii Studio 2.2.8 analysis software, with velocity and distance thresholds of 35 pixels/samples. Only fixations targeting the staff system and related to the actual reading of music notation were included in the analysis. Thus, fixations beyond a 45-pixel vertical distance from the staff were excluded. The limit was set exploratively with the goal of excluding clear outliers while including as many potentially task-relevant fixations as possible. With such a limit, in 100 % of the 101 analyzed trials, at least 70 % of the fixations occurring between the first and last MIDI note onsets fell within this visual area, and in 80.2 % of the trials, at least 90 % of them fell within the area. Fixations between the timestamps of the first note onsets of bar 1 (E4) and bar 8 (C1) were included, and it was required that the first note onset of bar 1 was, in all cases, the beginning of the actual performance. Equal-sized Areas of Interest (AOI) were created around each quarter-beat pattern, i.e. one individual quarter-note or a pattern of two eighth-notes, in order to assign each fixation to one quarter-beat pattern.

Third, the eye-movement and performance data were synchronized *post hoc*. Both of these separate recordings were launched by the experimenter on the metronome beats. For the key presses on each eye-movement recording that were intended to appear on the metronome beats at 2000 ms intervals (for making the cross on the screen appear and disappear – four such pairs for each of the participants), the mean of their temporal distance was 2012 ms (SD = 40 ms) and the 95 % confidence interval was (2004, 2020). (One outlier of the 101 trials with a temporal gap of 2447 ms was excluded from this analysis). To begin, for both recordings, the locations of the metronome beats were first estimated. For the performance data, the synchronization began by focusing on the final notes of each of the four trials for which no appreciable performance errors were observed, and on the experimenter's presses on the computer keyboard by which she changed the image on the computer screen, which was, likewise, to happen on the metronome beats. For each participant's MIDI recording (with the duration of one MIDI tick being 8.33 ms), the location of the metronome beat was estimated using the key presses representing the last notes (C4) of each trial in the 60 bpm tempo; considering that the motoric responses most likely have a minute delay to the actual beat, this involved taking the median of the decimal parts of the timestamps of the four key presses in question. By subtracting the resulting median from each timestamp of the respective MIDI recording, the estimated quarter-note beats were aligned with seconds in the MIDI event list. Next, focusing on the eye-movement data, all "key press" events, corresponding to changes of the image on the screen by the experimenter, were identified from the eye-movement recording. Given a set of 11 such key presses, the medians of the decimal parts of their timestamps were calculated. These represented the estimated beat in the eye-movement data. Finally, the two data sets were synchronized. Given that the MIDI recording was always launched by the experimenter simultaneously with the first key press with which she operated the appearance of the stimuli, the synchronization could be carried out by adding the time value of this first key press (with the decimal part corrected according to the estimated median) to each time value on the corrected list of MIDI time values. The synchronization was checked for errors by verifying that the durational spans between the key presses dictating appearance of the notated score and the first MIDI note onsets in the corresponding trials were 2 s; in two cases, this led to adjustments by 1 s. After the preparations, the data analysis proceeded to the examination of average fixation durations, eye-hand span and gaze activity (see Table 1).

3.3 Experiment III: Extracting information from music notation during silent reading

Experiment III was designed to examine the differences between adult novice and amateur musicians' silent reading of music notation, and the changes all participants manifested during a nine-month long music course. In addition, the experiment aimed to develop and test a categorization for analyzing the contents of verbal descriptions readers give of a musical score.

Participants. The 49 participants from Experiment I also participated in Experiment III. The participants were divided into groups of novices, less experienced (LE) amateurs and more experienced (ME) amateurs based on a written questionnaire administered before the first measurement. The group division based on self-reports was cross-checked by a sight-reading task conducted in the first measurement session. The sight-reading performances were analyzed for errors and temporal accuracy by two researchers collaboratively. The analysis did not suggest any changes to the group division.

Stimulus material. The notated melody for a 25-bar Russian folk song was prepared as the stimulus (see Appendix 1). The piece, adapted from an elementary school music book, is in Finnish, entitled "Punasaappaat", i.e. "Red boots". The melody was written with Finale music notation software, and the title and the lyrics were excluded.

Apparatus. Eye movements were recorded using a Tobii 1750 Eye Tracker (see section 3.1 above).

Procedure. Each participant took part in three different measurement sessions according to the protocol described in section 3.1 above. At the start of each measurement session, the participant was first introduced to the equipment – an electric piano and a computer screen behind it – and then asked to adjust the piano seat to a comfortable height, after which the eye tracker was calibrated. Subsequently, the participant was instructed to view the music notation appearing on the screen for 30 seconds and, immediately after its disappearance, to freely describe in speech what s/he had seen in it. The duration of the viewing period was chosen to accommodate reading the piece through in a tempo of bpm = 100 (without the repeat). The same instructions and stimulus were applied in all three measurements.

Data analysis. The analyses of eye movements were done with ClearView 2.7.1 analysis software developed by Tobii Technology AB. A fixation was defined to require that the gaze dwell within a 50-pixel radius for 60 ms or more. Eye-movement measures applied in experiment III are presented in Table 1.

All of the participants' verbal descriptions were scored for their *accuracy* and *integration* by a panel consisting of three of the members of the research group (the author, Erkki Huovinen, and Pekka Salonen). The two scores were kept separate so that, strictly speaking, only the accuracy score represented the "goodness" of description, whereas the intent of the integration score was to characterize the style of description (see Table 2). The accuracy score reflects the manner in which the descriptions succeed in correctly referring to non-trivial notational objects and situating them in a linear, temporal context. Each verbal description was assigned a score of 1 to 4 according to the highest matching criterion – one that correctly accounted for the use of musical terminology as well as the relationship between terminology and the piece of music as seen notated on the screen. The accuracy score is only concerned with the correct parts of the descriptions; even on the highest level, a description might also contain some details that are incorrect, if the criteria for the level are otherwise fulfilled. The integration score, on the other hand, reflects the manner in which the descriptions succeed in referring to relational entities or predicating properties of the whole piece of music. Each verbal description was assigned a score of 1 to 4 according to the highest matching criterion – one that correctly accounted for the use of musical terminology and/or the relationship between terminology and the piece of music. The integration score only concerns the quality of descriptions and not their accuracy; hence, not only might a richly integrated description contain incorrect details, but even the critical parts of such a description might sometimes be incorrect. Finally, two independent raters, both with a PhD in musicology, rated 30 (21 %) of the descriptions. Comparing the raters' ratings to the ones made by the panel, inter-rater reliabilities of 78 % and 87 % were reached.

Table 2. Accuracy and integration scoring criteria for the verbal descriptions in experiment III (*d* = description, *p* = piece of music seen on the screen).

	Score	Criteria
Accuracy	1	**Empty**: *d* does not clearly differentiate *p* from any number of other similar pieces. It may involve terms referring to comparatively trivial features of the notation system (e.g., 'there were sharps and flats', 'there were quarter notes'), or ones that do not require musical knowledge to be interpreted (e.g., the Finnish 'there were slurs' expressed using the common word for 'arc').
	2	**Discriminative**: *d* involves at least one feature that correctly refers to *p*, as opposed to other similar pieces. Such features have to be non-trivial considering the system of notation used, and they have to require music-specific knowledge to be interpreted. References to staff lines, the G clef, the time signature 2/4, and the flat sign in the beginning of the piece were considered trivial. Examples: symbol 'Gm' –> *d* = 'g minor chord'; symbol 'A' –> *d* = 'A section'. On this level, musical terminology is still not applied indexically or relationally: the elements are merely mentioned (e.g., 'There was a g minor chord', 'There was an A section'), but neither attached to particular locations in the score nor ordered with reference to each other.
	3	**Indexical**: *d* contains one non-trivial feature which correctly refers to *p* (as opposed to other similar pieces), and is applied indexically, that is, by attaching the elements mentioned to particular locations in the score (e.g., 'the *last* chord was d minor'). This category also accepted descriptions that contained incomplete formal descriptions of the whole structure of the piece (*e.g.*, 'ABA' or 'AAB' instead of 'AABA').
	4	**Richly accurate**: *d* contains several non-trivial features which correctly refer to *p* and are applied either relationally, ordering them with reference to each other (e.g., 'the B part was succeeded by another A part'), or indexically in conjunction with other indexical descriptions concerning the same musical parameter (e.g., 'The first chord was a d minor chord' *and* 'The B-part begun with a D7 harmony'). This category also accepted descriptions that contained complete formal descriptions of the whole structure of the piece (e.g. 'AABA').
Integration	1	**Enumerative**: *d* consists of references to individual musical or graphic elements, or to groups of them, but does not mention any interconnections between various types of elements. An enumerative approach dominates on this level (e.g., 'There was an A seventh chord, a d minor chord, an F…'). The elements mentioned may also be assigned locations, but this is done in spatial or visual terms (e.g., 'There was a G clef in the beginning of each staff system'). Mentioning the formal structure of *p* (*e.g.*, 'ABA', 'AABA') was also taken as an enumerative strategy.
	2	**Associative**: *d* refers to composite wholes involving connections between different types of elements (e.g., '*C* was *sharpened*' [note C, # sign], 'The *A part* was *repeated*' [symbol 'A', repeat sign]), or *d* predicates an abstract property of some element, other than temporal location (e.g., 'There were many *high* notes'). However, rather than describing *p* as a whole, *d* only mentions objects that are parts of it.
	3	**Integrative**: *d* predicates one non-trivial concept or idea of *p* as a whole. This might involve statements of key or registral range, or statements concerning the overall mood of the piece, its relative difficulty, etc.
	4	**Richly integrative**: *d* predicates several different non-trivial concepts or ideas of *p* as a whole. As such, all of these statements might already be found on level 3, but the difference is that now *p* has been described from several different perspectives (e.g., 'The piece was in d minor and the melody was easy').

4. OVERVIEW OF THE EMPIRICAL STUDIES

This dissertation consists of five empirical studies. Studies 1 and 2 were based on Experiment I, studies 3 and 4 on Experiment II, and study 5 on Experiment III.

1) **Penttinen, M., & Huovinen, E. (2009). The effects of melodic grouping and meter on eye movements during simple sight-reading tasks. In J. Louhivuori, T. Eerola, S. Saarikallio, T. Himberg, & P. Eerola (Eds.), *Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music*, Jyväskylä, Finland. (pp. 416–424).**

The main aim of this exploration was to begin a series of systematic eye-tracking studies focusing on basic notational and music-theoretical factors. The study addressed the importance of temporal control in eye-movements research on music reading, and was among the first on musical sight-reading to use eye-movement measures similar to those applied in several text-reading studies. The main research topics addressed (i) performance and eye-movement indicators of skill and skill development and (ii) the effects of melodic grouping and metrical structure on eye movements during sight-reading. Due to the exploratory nature of this orienting study, several more specific research questions were created around the main topics.

The original number of 49 participants from Experiment I (see section 3.1) was cut down by missing eye-movement or performance data and finally, 32 participants (future elementary school teachers) were included in the analyses. The participants were divided into three groups according to their musical background: novices (*n* = 12), less experienced (LE) amateurs (*n* = 11) and more experienced (ME) amateurs (*n* = 9). The participants attended two measurement sessions during their music training period, the first of which took place at the beginning of the music training, and the second after three months of musical tuition. In each measurement, a participant performed four simple sight-reading tasks consisting of five-bar melodies in C major (see Appendix 1). The following eye-movement measures were calculated for each note as a percentage of a participant's overall fixation time: total fixation time, first-pass fixation time, and look-back time. For research topic (i) all participants' performances were analyzed, and for research topic (ii) all temporally accurate performances (*n* = 189) were examined.

The results demonstrated, first of all, a monotonous improvement in the accuracy and timing of the performances from novices to LE amateurs to ME amateurs. ME amateurs generally had a higher percentage of look-back time than novices, suggesting that novices sight-read in a more stepwise fashion, while more experienced amateurs perform more back-and-forth eye movements during sight-reading. No significant indicators of skill development between the two measurement sessions emerged for any of the participant groups. Second, when examining all temporally accurate performances and combining the two measurements, first notes of bars were fixated less and second notes of melodic groups were fixated more than other notes. When large melodic intervals coincided with bar lines, re-inspections to individual notes tended to diminish as sight-reading proceeded; when the two large intervals were located within bars, on the other hand, there was a tendency for look-back behaviour to remain similar throughout the melody. In larger descending intervals, first notes received more first-pass fixation time than latter notes. In total, then, several issues worth a closer investigation manifested themselves on the level of eye movements, emerging both in relation to sight-reading skill as well as the reading of a simple musical score in a temporally controlled condition.

2) **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.**

This study was a step forward from the previous orienting study and examined skill development in sight-reading across a longer training period. The main issue of interest was to assess the indicators of skill development in both (i) musical performances and, most importantly, (ii) in eye-movement behaviour of adult beginners in simple sight-reading tasks during and after a 9-month music training period. Concerning the latter topic, the focus was on the role of bar lines as visual organizers in the reading of music notation as well as on the note-identification strategies of the novice participants.

The original number of 49 participants (future elementary school teachers) from Experiment I (see section 3.1) was cut down by missing eye-movement data in one or more of the total of 12 performed melodies. Thus, in the study, 15 novices' development was assessed in three measurements at different time points and compared to 15 amateur musicians, who reported to understand music notation and had had musical training prior to the study (combining here the groups of less and more experienced amateurs of Experiment I). The participants attended three laboratory sessions and performed four simple sight-reading tasks in each session.

During the training period, the novice musicians' sight-reading performances improved regarding their overall temporal accuracy and performing of larger melodic intervals. The amateurs, as expected, produced temporally accurate, correct performances from the first measurement onward. In terms of eye movements, the novices' allocation of fixation time within metrical units gradually approached a pattern demonstrated by the amateurs, showing increased sensitivity to metrical divisions. The pattern was characterized by larger average total fixation times on the latter halves of the bars (each bar containing four quarter notes). Concerning larger melodic skips in otherwise stepwise melodic contexts, an analysis of fixation times suggested that the novices' visual processing of skips did not proceed in terms of note comparison across the skip but rather through a direct identification of the notational symbols involved. To sum up, the novices' early indicators of skill development in music reading during the nine-month long training period were manifested in the increase of temporal stability and performance accuracy, re-allocation of fixation time across metrical units, and the faster identification of individual note symbols.

3) **Penttinen, M., Huovinen, E., & Ylitalo, A. (2012). Unexpected melodic events during music reading: Exploring the eye-movement approach. In E. Cambouropoulos, C. Tsougras, P. Mavromatis, & K. Pastiadis (Eds.), *Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Conference of the European Society for the Cognitive Sciences of Music,* Thessaloniki, Greece. (pp. 792–798).**

A series of two exploratory studies, one pilot and one experiment proper, examined how melodic alterations affect eye movements during a temporally controlled task of rehearsed reading. The methodological goal was to explore how such phenomena might be addressed through eye-tracking methodology. The pilot experiment tested a research design suitable for examining the eye-movement effects of unexpected melodic events during reading and performing a rehearsed, familiar melody. In addition, the pilot study explored the preliminary hypothesis that in a temporally controlled performance, unexpected melodic events cause deviations from the average course of reading. The aim of the main experiment was to statistically examine how unexpected melodic deviations from a familiar melody affect the allocation of fixation time and saccadic eye movements for musically experienced young adults.

For the pilot experiment, five experienced amateur musicians also participating in Experiment I (see section 3.1) were chosen based on their (i) extensive musical training and music reading ability, (ii) successful eye-movement and performance recordings in all three measurements for the three melodies in question and (iii) temporal and melodic accuracy of the to-be-analyzed performances. The participants performed three variations of a familiar tune, "Mary Had a Little Lamb", in time with a metronome. The findings showed that the alterations from the familiar melody did *not* increase the relative total fixation time for the altered bars. Instead, there was a relative increase of incoming saccades to the target bar and the bar preceding it, suggesting that the strategy for overcoming more challenging musical material is related to more frequent visits to the problematic area instead of directly allocating more fixation time to it. Thus, it appears that unexpected alterations within a familiar melody did result in re-adjustment of the visual processing. Most importantly, however, the pilot experiment revealed several issues for improvement in the experimental design. The design was accordingly modified for the main experiment.

The 34 participants of the main experiment (Experiment II; see section 3.2) were music education students (future elementary school teachers; n = 21) and students of music performance (n = 13). The participants performed the original version of "Mary Had a Little Lamb" and two of the variations created for the pilot experiment. The melodic alterations were found to affect both the relative fixation times and the distribution of incoming saccades. First, similarly to the pilot study, there was no increase of relative fixation time at the unexpected bars themselves. Instead, the melodic alteration had the effect of decreasing the relative fixation time allocated to the immediately following bar. Second, the altered bars generally had the effect of increasing the proportion of incoming saccades to the bars before them, this result also being partly aligned with the pilot experiment. In addition, the immediately following bars showed a decrease in the relative amount of incoming saccades. These preliminary findings propose that problematic areas in music notation during temporally controlled rehearsed reading are not simply fixated more than other areas. This is contradictory to findings in text-reading research, where problematic words tend to collect more fixation time. To further clarify the issue, performance and eye-movement data should be synchronized and examined in more detail. However, these experiments do demonstrate that eye-tracking methodology allows us to tap even such aspects of the musical performances which do not result in clear audible effects, say, in the form of mistakes or expressive deviations. In addition, it seems that even seemingly trivial modifications of simple notated music affect the course of the visual processing even for skilled performers.

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

This study examined how melodic alterations affect skilled performers' eye movements during temporally stable, accurate performances of a simple melody. The study focused on two measures: the eye-hand span, i.e. the distance between the fixated note and the performed one, and gaze activity, measured as the number of inspected quarter-beats between two beat onsets. The study aimed to complement prior findings on background-related differences in the eye-hand span by controlling the performance tempo, and introduced gaze activity as a new eye-movement measure in music-reading studies.

The 38 participants were music education students (future elementary school teachers; *n* = 24) and students of music performance at a Finnish arts academy or conservatory (*n* = 14; see Experiment II, section 3.2). The musical stimuli were the original version of "Mary Had a Little Lamb" (in C major) as well as two of its variations, where one of the bars had been slightly altered. The participants performed all melodies in the meter of 2/4 in time with a metronome set at 60 bpm. Only temporally stable and flawless performances were analyzed.

All participants typically operated with eye-hand spans of 1 quarter-beat or less, the result supporting previous notions on the average temporal length (1 s) of the eye-hand span. During the performance of one beat, the performers most often inspected two quarter-beat areas. However, the performance majors operated with shorter average fixation durations and demonstrated a higher frequency of (i) eye-hand spans longer than 2 quarter-beats and (ii) gaze activity of more than 2 quarter-beats, compared to the music education students. Thus, increase in musical experience increased the processing efficiency, despite the simplicity of the tasks and similarity in audible outcomes. However, the one-bar long melodic alterations affected the applied measures so that the general effects of musical experience disappeared in the face of the alterations. As an additional finding, simple rhythmic patterns, namely individual quarter-notes versus patterns of two eighth-notes, also affected the examined measures: shorter eye-hand spans and increased gaze activity appeared more often, while performing patterns of two eighth-notes. In total, then, the study was able to demonstrate that even experienced musicians' eye movements are affected by minor alterations in a familiar melody, as well as some almost trivially simple, minute music-structural details. The findings highlight both the flexibility and limitations of the eye-hand span in temporally controlled, accurate performances.

5) Penttinen, M., Huovinen, E., & Ylitalo, A. (in press). Silent music reading: Amateur musicians' visual processing and descriptive skill. *Musicæ Scientiæ*.

This longitudinal study examined the extraction of information from music notation in a non-performance task. The focus was on the effects of prior musical experience on silent music reading and on the changes adults with varying musical background demonstrate during a nine-month long music course. In addition, the study explored, irrespective of the participants' musical background, potential styles of silent reading occurring in the context of a freely-defined music-reading task.

The original number of 49 participants from Experiment III (see section 3.3) was cut down by missing data and finally, data from 37 future elementary school teachers were included in the analyses. The participants were divided into three groups according to their musical background: novices (n = 16), less experienced (LE) amateurs (n = 11) and more experienced (ME) amateurs (n = 10). The participants took part in three measurement sessions during a compulsory music course designed to prepare even the musical novices to teach the subject in elementary school. In each measurement, a participant was instructed to look at a musical score from a computer screen for 30 s and, at its disappearance, describe in speech what s/he had seen in the image.

In general, greater musical experience led to more accurate and integrative verbal descriptions of the music at hand and more linear scanning of the notated music. Also the average fixation duration tended to shorten along with increasing musical experience. For all participants, the average saccade length as well as the accuracy and integrative scores increased during the training period. Importantly, however, no interaction between musical background, time and any of the measures was observed; thus, the novices, with the most to gain from the musical training, did not demonstrate any greater changes across time in their descriptions or eye-movement behaviour than the groups of amateur participants. Finally, a cluster analysis revealed three silent-reading styles: *elementary processors'* descriptions lacked both accuracy and integrative quality; *accurate analyzers* produced factually correct descriptions which were still low in their integrative quality; and *accurate integrators* achieved high scores on both accuracy and integration in addition to a tendency of operating with the shortest average fixation durations (the latter effect was observed only in the first measurement). In general, the silent-reading task was suggested beneficial for studies on music reading for the following reasons. First, the task allows the assessment of even novices' comprehension of music notation, despite their lack of performance skills. Second, examining experts' silent-reading strategies could enlighten the understanding of expertise in music reading. Third, the task of silent reading is a useful bridge between research on music reading and research in other non-textual domains, as the relaxing of the motoric demands allows music notation to be approached from several different perspectives and under various types of task requirements.

5. MAIN FINDINGS AND DISCUSSION

This dissertation investigated how skill and skill development in music reading are reflected by various eye-movement measures, and developed and tested research settings by, for example, including carefully created stimuli, temporal control of performance tasks and alternative ways to assess the comprehension of music notation into the research designs. The following sections address these main aims.

5.1 The reflection of skill and skill development in eye-movement measures

It has been stated that increase in musical experience is connected to the use of shorter average fixation durations during music reading (Goolsby, 1994a; Truitt et al., 1997; though note the lack of similar finding of Gilman & Underwood, 2003). However, these findings have been obtained in research conditions where the performance tempi or musical stimuli were not controlled for. The present experiments confirm that also in temporally commensurable performances, a similar pattern emerges (see Table 3). In Experiment II, during the rehearsed performance of a familiar melody, music performance students indeed operated with shorter average fixation durations than music education students, though all participants produced errorless and temporally stable performances. Note that the effect was observed despite the fact that the musical stimulus, a children's song, would at first glance seem almost trivially simple for all of these musically experienced performers. Thus, the experienced performers' more efficient processing of musical symbols appears even in a performance situation where such efficiency would not be a necessity. In addition, a tendency toward a similar mark of processing efficiency was noted during the task of silent reading (Experiment III). This finding repeats the observation of Waters and Underwood (1998), although in remarkably different study conditions. After this general observation, the main findings related to the eye-movement measures presented in Table 1 are discussed, separating here visual processing during instrumental performance (sight-reading and rehearsed reading) and silent reading with no performance requirements.

5.1.1 Sight-reading and rehearsed reading

In Experiment I, measures traditionally applied in text-reading research were exploited in the context of music reading. The differentiation between *first-pass fixation time* and *look-back time* allowed the separation of the first encounter of a musical symbol from its later inspections. Especially in temporally controlled *prima vista* performances, when the performer is left without the benefits of prior encounters with the musical material, the first encounters with musical symbols are of particular interest. Though some of the information is processed within the perceptual span and thus without directly fixating the symbol, these first inspections should reflect the effort needed for successful processing of the symbol (or group of symbols). In addition, the back-and-forth eye movements observed by Goolsby (1994b) can quite easily be explored from a data set by simply calculating the amount or proportion of look-back fixations. Thus, during sight-reading, it is useful to focus on the different roles of the first-pass and look-back fixations.

First, it was observed that novice musicians' relative first-pass fixation time allocated for latter notes of larger intervals shortened with training (Table 3; see also Appendix 1). These notes required the recognition of the pitch instead of merely proceeding along the score in a simple stepwise, up and down fashion. The analysis necessarily consists of the effects of improvements in reading skills as well as performance ability, but one may yet assume that also an increase in the efficiency to recognize individual notes was observed through this measure. On the one hand, the observation is aligned with the notion that average fixation duration is an indicator of sight-reading skill. On the other, the analysis was also performed only for temporally commensurable performances, which eliminates the effects of performance duration. Second, more experienced amateurs spent more time on look-backs than novices, suggesting that during the reading of previously unknown melodies, novice performers rely more heavily on first-pass processing, whereas more experienced performers, generally operating with shorter fixation durations, perform more back-and-forth eye movements (cp. Goolsby, 1994a, 1994b). We need to note here, however, that also the amateur performers spent a great majority of their total fixation time (75–80 %) on first-pass fixations.

The exact purpose of the look-back behaviour can only be explored by synchronizing the performance and eye-movement data and examining *when* these re-readings occur. This was carried out in Experiment II, this time in the context of rehearsed reading. The general observation on the course of visual processing during temporally controlled performances was that the gaze was generally one second or less ahead of the point of performance during a beat onset. The finding is aligned with prior observations, where the data was collected without the exact control of the performance tempi (Furneaux & Land, 1999; Wurtz et al., 2009). In addition, usually two (or more) quarter-beat areas were inspected during the performance of one quarter-beat (cp. Truitt et al., 1997). In 97.5 % of the cases, gaze activity was targeting music *ahead* of the point of performance. The findings thus suggest that the eye-hand span indeed has temporal limitations, and the gap between the point of gaze and the point of performance is moderate at most. However, the shortness of the eye-hand span is compensated by inspecting a slightly wider area of the score between beat onsets, as suggested by the performers' gaze activity. It thus seems that more experienced performers' increase in look-back time observed in the context of simple sight-reading tasks could be due to the fact that individual note symbols are first inspected *prior* to their performance. The notes thus gain little first-pass fixation time, while the look-backs to these particular symbols are from the time of the actual performance of the notes. Thus, similarities in the course of the visual processing during temporally controlled sight-reading and rehearsed reading seem to appear, as far as fixation durations and re-inspections are concerned.

As noted above, and also suggested by Truitt and colleagues (1997), the performer's gaze generally seems to remain very close to the point of performance. On the one hand, one could have assumed that the simplicity of the melodies in Experiment II would have allowed the performers to read well ahead in musical time, instead of focusing on the currently executed note (or the following one) during a note onset. Apparently, this was not the case. On the other hand, one could assume that the simplicity of the tasks decreased the need to perform any extensive eye movements, and that the situation would be different when performing more complex music. However, the findings here are in line with the observations of Furneaux and Land (1999) and Wurtz and colleagues (2009), who applied more complex musical stimuli. In addition, in such a case, one would also expect to find no background-related differences between the two groups of skilled participants, as even the less experienced performers had no difficulties in performing the tasks. However, the future professional musicians, compared to students of music education, operated significantly more often with eye-hand spans of more than one second, while also demonstrating more gaze activity of more than two beats and operating with shorter fixation durations (Table 3). Overall, then, in these simple tasks with familiar musical stimuli (only disturbed by the total of two altered bars in the three performed melodies) performed in relaxed tempo, the more experienced performers did *not* act with minimal processing efficiency, but applied reading strategies which differentiated them from performers with less musical experience.

Table 3. Skill and skill development in eye movements: the main findings of studies 1–5. Means and standard deviations.

			Novices	**Less experienced amateurs**	**More experienced amateurs**	**Music education students**	**Performance students**
Sight-reading	First-pass fixation time for the 2nd note of a large interval (%)[a]	*Before training*	8.2 (4.5)				
		After 9 months	6.3 (4.6)				
	Look-back time (%)[b]	*Before training*	11.6 (5.3)		19.0 (10.9)		
		After 3 months	14.3 (8.9)		24.1 (9.0)		
Rehearsed reading	Eye-hand spans of more than 1 s (%)[c]					7.8	14.3
	Gaze activity of more than 2 beats (%)[d]					28.0	36.6
	Fixation duration (ms)[e]					640 (180)[f]	520 (155)[f]
Silent reading		*Before training*	370 (90)	360 (80)	310 (100)		
		After 9 months	370 (100)	350 (111)	280 (60)		
	Saccade length (pixels)[g]	*Before training*	161 (39)	158 (29)	173 (49)		
		After 9 months	163 (36)	177 (38)	188 (41)		
	Linear scanning (%)[h]	*Before training*	39.3 (10.6)	44.3 (10.2)	47.5 (14.4)		
		After 9 months	36.9 (8.7)	39.2 (8.8)	39.9 (9.4)		

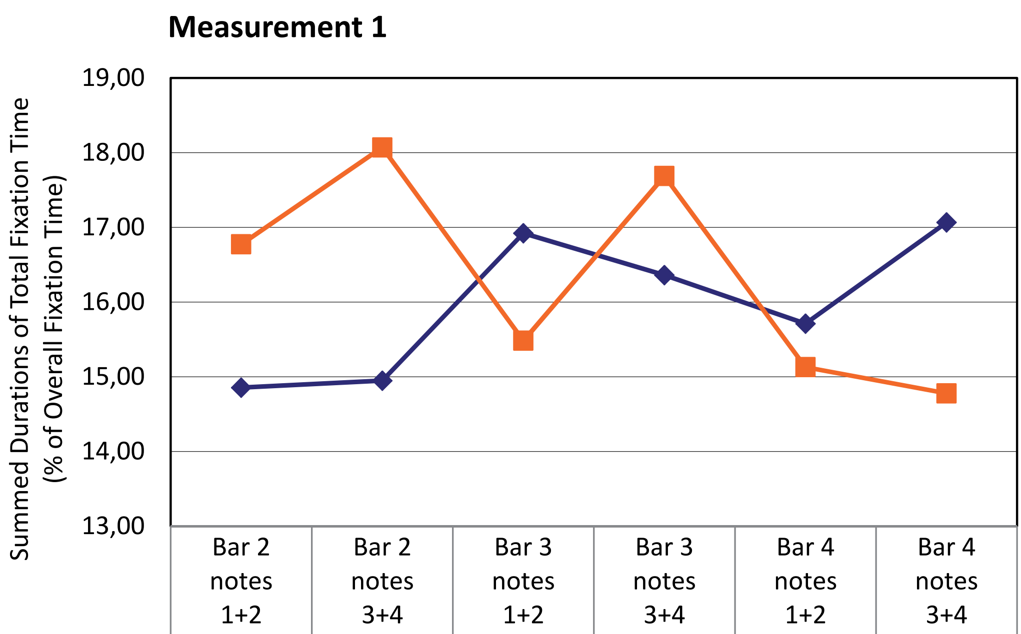
[a] Proportion of overall fixation time. Average values for temporally accurate performances. Sign. difference between measurements.
[b] Proportion of overall fixation time. Sign. between-group differences in both measurements.
[c] Sign. between-group difference in distributions.
[d] Sign. between-group difference in distributions.
[e] Rehearsed reading: sign. between-group difference. Silent reading: near-sign. between-subject effect of participant group.
[f] Average values for the original versions of "Mary Had a Little Lamb".
[g] Sign. within-subject effect of measurement.
[h] Sign. between-subject effect of participant group.

Incongruent endings of melodies have been noted to result in the use of more fixations and longer average fixation durations (Ahken et al., 2012). Now, the question still remains whether the eye-hand span is affected by unexpected musical material during a temporally controlled performance. In Experiment II, the presence of unexpected melodic alterations was assumed to present the performers a problem-solving task, as they needed to overcome the violation of their melodic expectations in order to perform accurately and in the given tempo. Both music education and music performance students managed to perform the alterations with only a minute amount of error, but interestingly, background-related differences in the visual processing nevertheless appeared. Performance students used larger eye-hand spans and greater gaze activity more frequently than music education students, but this difference disappeared in the altered bars. Thus, in a performance task that quite obviously did not present any excessive challenges to any of the participants, a simple violation of the performers' melodic expectations was immediately reflected in the visual processing: the more experienced performers' eye-hand spans shortened and gaze activity decreased to a notable extent. Thus, these more experienced performers were applying their automated, more efficient reading styles even in these simple tasks, but the melodic alterations disrupted the process. Whether the effects are caused by difficulties in motor programming, evaluating the correctness of the performance (as it violated the performers' expectations), or any other reason, the present experiment cannot answer. However, the findings demonstrate that in order to examine all of the phases of music reading during a performance from recognition of the symbols to temporally controlled execution, even simple performance tasks would suffice.

The present experiments also addressed the relationships between eye movements, musical experience and some basic features of music notation. To begin, in Experiment II, the research design allowed the comparison of eye-hand span and gaze activity during the performing of quarter notes or (even or dotted) eighth-note patterns (see Appendix 1). Irrespective of musical background, shorter eye-hand spans and increase in gaze activity appeared at the onsets of eighth-note patterns, compared to quarter-notes. Thus, even such a seemingly trivial fact as whether a beat in a familiar, simple melody contained one or two note symbols had effects on the course of the visual processing. As the eye-hand span was more often shorter for the more experienced performers during the performance of melodic alterations, the eye-hand span seems to be highly sensitive to the contents of the musical stimuli. The already relatively short distance between the points of performance and gaze can thus quite easily become even shorter. The everyday conception of high-level music-reading (or sight-reading) ability being marked with a large distance between the hand and the gaze did not find support in any of the findings of Experiment II. On the contrary, even elementary details were found to decrease the eye-hand span, and even so for skilled performers.

Yet another finding related to basic features of music notation was observed in Experiment II: when eighth-note patterns were located at the ends of metrical units (i.e. on the latter beats of 2/4 bars), the skill-related differences in the length of the eye-hand span disappeared. In Experiment I, when examining the distribution of total fixation time in sight-reading simple melodies, it was observed that amateur musicians tended to fixate relatively more on latter halves of bars (Figure 2). After nine months of training, and simultaneously with improvement in performance ability, novice musicians approached this pattern of allocation of fixation time. In the final measurement, the initial significant difference in the patterns of fixation time between the groups disappeared. Overall, the findings support the notions of Drake and Palmer (2000) and Sloboda (1978) that the highly typical feature of Western music notation, namely the visual dividing of scores into bars by vertical bar lines, acts as an organizer of visual processing. It seems that in performances performed in relatively slow tempi, the latter halves of bars collect more fixation time. For experienced performers, the eye-hand span may also shorten. More research on this topic is needed, but the findings here do demonstrate that even relatively basic features of music notation affect the music-reading process. Together with the above-mentioned findings on differences in processing different rhythmical patterns and the findings by Ahken and colleagues (2012), this discovery also denotes that even in high levels of performance ability, performers' visual processing is, to a notable extent, affected by the low-level visual features of music notation.

To summarize, the findings suggest the following. First, we need to note that in all the above-mentioned performance tasks, the tempo was set at 60 bpm. In the sight-reading tasks, the melodies consisted almost only of quarter notes; in the task of rehearsed reading, the melodies contained quarter notes and eighth-note patterns. To begin, experience in music reading is related to the use of shorter fixations even when the performance tempo is controlled for (and even in non-performance tasks; see section 5.1.2 below). Novices tend to sight-read in a more stepwise fashion, and their skill development seems to be apparent in the shortening of first-pass fixation time for more challenging note symbols. More experienced performers, using shorter fixation durations, are able to inspect more of a musical score during the performance of one beat, though the gaze generally tends



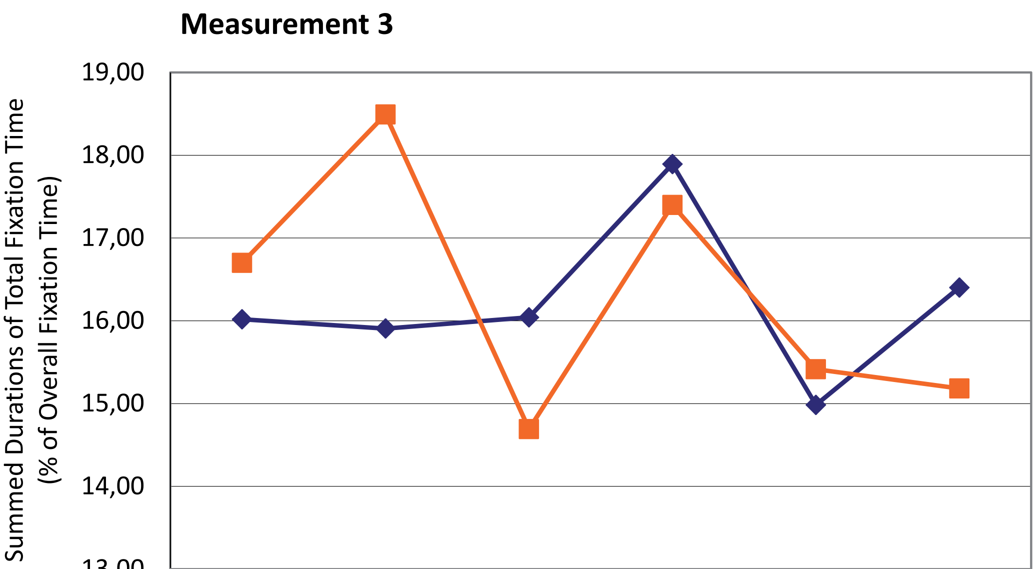
	Bar 2 notes 1+2	Bar 2 notes 3+4	Bar 3 notes 1+2	Bar 3 notes 3+4	Bar 4 notes 1+2	Bar 4 notes 3+4
Novices	14,85	14,95	16,92	16,36	15,71	17,07
Amateurs	16,77	18,07	15,48	17,69	15,13	14,78

Figure 2. The summed durations of total fixation times for the first and last two notes in bars for novices and amateurs in study 2 before training (p. 38) and after nine months of training (p. 39).

to be very close to the point of performance during the onset of a musical beat. Thus, the typical strategy of a skilled performer seems to be to inspect upcoming musical material and then return towards the point of performance during the actual performance of the note or the note pattern. These eye-movement patterns appear similar both during sight-reading and rehearsed reading. The eye-hand span has, as suggested before, a temporal average of around one second, but with increasing expertise, performers are able to use the maximal eye-hand span more often. These skill-related findings are, however, only half the truth. Apart from these "top-down" processes, also the "bottom-up" effects of specific features of music notation, such as pitch heights, metrical division, rhythmic patterns and unexpected melodic events, should be investigated further. According to the findings, these features of music notation apparently play as significant a role in organizing the visual processing during music reading as do, for example, familiarity or length of words in text reading. Most importantly, these elementary details affect all music readers from beginners to experts.

5.1.2 Silent reading

Experiment III focused on the extraction of information from music notation in a non-performance task. The viewing process was without specific instructions (only asking the participant to "describe what s/he had seen") to allow the viewers to express their understanding of the written music with respect to their individual interpretations of the task requirements. Overall, the intuitive approaches to a silent-reading task of readers with varying musical backgrounds were manifested in both how the score is *read* as well



	Bar 2 notes 1+2	Bar 2 notes 3+4	Bar 3 notes 1+2	Bar 3 notes 3+4	Bar 4 notes 1+2	Bar 4 notes 3+4
Novices	16,02	15,91	16,04	17,89	14,98	16,40
Amateurs	16,70	18,49	14,69	17,40	15,41	15,18

as how it is *described*. To begin, as noted above, the more experienced amateurs tended to make use of shorter average fixation durations. In addition, they scanned the score in a more linear manner than novices (Table 3). It thus seems that music-reading skill affects the way music notation is inspected, when given no direct instructions on what to look at or search for (cp., e.g., the specific task definitions of Gilman & Underwood, 2003). Contrary to expectations, however, the novices did not demonstrate significant changes during the nine-month training in average fixation duration nor in linear scanning. Thus, the more experienced amateurs' "reading approach" required more familiarity with music notation and/or performance experience than the novice participants had time to gain during their training. Instead, an increase in the average length of saccades was observed, but this effect was similar to all participants. This observation cannot therefore be interpreted simply as an indicator of developing skill; more likely it reflects a learning effect for this specific viewing task (see section 5.2 for further discussion on the research setting).

Considering the qualitative aspects of the extracted information, prior musical experience was, as expected, connected to what might be called more sophisticated descriptions of notated music: with increase in experience, the participants were not only able to use relevant terminology more *accurately*, applying designations that correctly referred to the properties of the particular piece of music, but they also produced more *integrative* descriptions. In other words, they referred more often to global aspects of the melody such as key signature, mood, or experienced simplicity, as opposed to a piecemeal listing of separate visual items. Contrary to expectations, the novice music readers did not demonstrate any intuitively integrative understandings of the musical image. An explanation here might be that the novices, despite the openness of the silent-reading task, interpreted it to be one of music analysis and, being aware of their lack of technical concepts, were reluctant to express their broader intuitive interpretations by folk terminology. All participant groups improved their average scores for both semantic categories across the measurements. Surprisingly, however, the novices, with the most to gain from the training, did not improve their scores any more than the two other groups did.

In such freely defined visual tasks, musical background may not be the only factor defining how a musical score is read. In order to explore whether the data would demonstrate any styles of silent reading irrespective of readers' musical background, the eye-movement and verbal data were explored together and three styles of reading were discovered (Table 4). *Elementary processors'* descriptions lacked both accuracy and integrative quality, while *accurate analyzers* produced factually correct descriptions that were still relatively low in their integrative quality. *Accurate integrators*, on the other hand, achieved higher scores on both accuracy and integration in addition to demonstrating a tendency for shorter average fixation durations (the last characteristic appearing in the first measurement) than the two other groups. The distinction between the two last-mentioned groups indicates that the semantic categorizations did in fact manage to tap separate semantic dimensions of the descriptions, as intended. Presumably these different silent-reading styles reflect differences in the viewers' interpretations on the task requirements; accurate analyzers, for instance, may have considered the task to be a memory test, and answered and processed the score according to this interpretation. In comparison between measurements 1 and 3, 51 % of all participants remained faithful to their initial silent-reading style also in the last measurement, whereas 35 %

of the participants moved to a style characterized by higher verbal scores in the final measurement. The cluster of accurate integrators, dominated by members of the group of more experienced amateurs in the initial measurement, contained members of all participant groups in the final measurement.

Overall, then, musical background influences the reading of music also in silent-reading situations. This is manifested in the average fixation durations as well as in linear scanning of the music. However, in addition to musical experience, the viewers' interpretations of the task requirements are also reflected in their verbal descriptions of written music, and here the effects of musical background are confounded. The exact relationship between musical background and other personal features affecting the silent-reading process is a task for future studies. However, the present dissertation suggests that such explorations should indeed be administered in order to better understand the learning to comprehend music notation and to explore the multiple uses of notated music.

Table 4. Descriptions of the three silent-reading styles: measurements 1 and 3. Means and standard deviations.

		Elementary Processors	**Accurate Analyzers**	**Accurate Integrators**
Measurement 1	Average score for accuracy	1.00 (0.00)	2.60 (0.74)	2.86 (0.90)
	Average score for integration	1.00 (0.00)	1.93 (1.03)	3.14 (1.07)
	Average fixation duration (ms)	365 (97)	368 (86)	278 (36)
	Novices	10	6	0
	Less experienced amateurs	4	5	2
	More experienced amateurs	1	4	5
	n	*15*	*15*	*7*
Measurement 3	Average score for accuracy	1.45 (0.52)	3.45 (0.52)	3.07 (0.80)
	Average score for integration	1.00 (0.00)	1.91 (0.30)	3.40 (0.51)
	Average fixation duration (ms)	340 (90)	378 (104)	313 (97)
	Novices	8	5	3
	Less experienced amateurs	3	3	5
	More experienced amateurs	0	3	7
	n	*11*	*11*	*15*

Note. Scoring scale for accuracy and integration from 1 to 4.

5.2 Review of the applied research settings

The second aim of this dissertation was to systematically examine eye movements during sight-reading, rehearsed reading and silent reading, and to develop and test research settings suitable for such explorations. The generalizability of prior studies on eye movements in music reading has suffered, to some extent, from the variability in the use of stimuli, the smallish numbers of participants, and the lack of temporal control in performance tasks. Next, the three experiments presented in this dissertation are discussed both in terms of their strengths and weaknesses, and suggestions for future studies are presented.

In Experiment I, two major methodological benefits were implemented: the creation of well-structured sight-reading tasks and temporal control of the performances. First, as the stimulus melodies consisted almost solely of quarter notes, the effects of larger melodic intervals were clearly distinguishable from effects caused by, for example, different rhythmic patterns. The creation of simple melodies and their inversions produced pairs of melodies as similarly structured as seemed possible. This strategy for creating stimulus melodies is directly applicable to studies on the eye-movement effects of multiple types of melodic or rhythmical patterns. Overall, similar interpretations of the findings would have been difficult to make, had the stimuli been “authentic”, i.e. originally composed for performance purposes. The second benefit, temporal control, introduced to the setting the all-important time frame that has lacked in almost all of the prior studies on eye movements in music reading. All performers needed to adjust their visual processing within the given time frame, and this resulted in a notable amount of temporally comparable performances. Note also that some of the novices were able to operate within the given temporal framework, and these performances allowed the close examination of the eye-movement indicators of skill development without the confounding effect of differences in performance durations. The melodies were simple and the performance tempo relaxed enough to bring forth the concurrent development of the skills of sight-reading and performance ability of the novice participants. Finally, it needs to be noted that the research setting seemed quite convenient for the participants, as 46 of the 49 participants attended all three sessions during the nine-month time interval, with somewhat minimal compensation. The setting thus appeared not to be too strenuous even for the novice participants, for whom the task was, without a doubt, demanding.

However, possible improvements on the design also became apparent and should be addressed in future studies. Despite the simplicity of the melodies, many novices were still struggling at the first measurement, and produced erroneous and temporally unstable performances. Thus, it seems that even simpler melodies should have been created, if to no other purpose than to offer the novices as easy as possible a start for a series of performances. Furthermore, a precise analysis of the eye-movements effects of large melodic intervals in temporally controlled performances should yet be confirmed by also including in the series of melodies stepwise ones, i.e. melodies lacking the larger intervals. (In a follow-up experiment not included in this dissertation, the research setting was indeed complemented by adding such melodies together with an additional performance tempo, this time performed by experienced musicians.) Finally, the use of an “authentic” style of writing the melodies led to the slightly uneven sizes of the five bars of the melodies, and this caused minor difficulties for the analyses. (This matter also was addressed in the above-mentioned follow-up experiment.)

In Experiment II[1], musically experienced participants performed a familiar melody and its variations in a given tempo. The aim was to collect a large sample of errorless, temporally stable performances, which required the inclusion of study participants with extensive musical background. First, creation of stimuli with visually non-salient melodic alterations appeared to be successful, despite the original melody's obvious simplicity with regard to the musical experience of the participants. Presenting the performers with a familiar melody and allowing them time to practice eased the interpretation of the results. Indeed, the observed eye-movements effects are most probably due to the slight deviations from the original melody, as the rhythmic characteristics of the melody were not altered. Second, it is remarkable how even such a seemingly trivial task brought about obvious between-group differences relating to musical experience, as well as differences in the processing of extremely simple rhythmic patterns. On these grounds, future studies should also apply systematically created musical stimuli, controlling as many aspects of the written music as possible. Finally, the loss of data was remarkably low due to the level of experience of the performers, simplicity of the task, and, to some extent, the use of an upgraded eye-tracking device.

Although the current design allowed comparison between visual processing of one highly familiar melody and its variations, the inclusion of a number of other simple melodies created in a similar style would have strengthened the results. However, should one prefer to use authentic melodies, it first needs to be explored which melodies would be somewhat equally familiar to particular participant groups. Libraries of suitable melodies could perhaps be created with collaboration between researchers and music instructors, at least within specific cultural contexts. Another option would be to compose completely new melodies by carefully controlling their contents, and allow enough time for the participants to familiarize themselves with the stimuli before the actual experiment (standardizing the level of performance accuracy, after which the actual reading task could begin). All in all, slightly different issues come to the fore, dependent on whether one is designing research settings for a study on sight-reading or on rehearsed reading. The latter type of music reading, especially, is necessarily confounded by multiple intervening factors, and the roles of all of them are beyond the reach of one individual study. To minimize the amount of such factors, the musical stimuli and study conditions should be maximally controlled. It is suggested here that the distinction between these two types of reading music during a performance should be the starting point for creating research hypotheses – the first step in wondering how to examine or measure the issue of interest.

Experiment III addressed the topic of silent reading, a type of music reading often neglected or considered as a part of the sight-reading procedure. Despite the exploratory nature of the experiment, it nevertheless managed to expose differences between participants of varying musical experience. The experiment also demonstrated one way of examining the extraction of musical information during a non-performance music-reading task. The careful creation of the categorization of retrospective verbal descriptions was intended as an opening for similar methodologically divergent approaches for understanding the process of information extraction. In general, after this first exploration, one sees a multitude of options for varying the setting by altering the duration of the viewing process, adding the number of stimuli, or presenting the viewers with more specific task demands.

1 This section only discusses the experiment proper of Experiment II. For discussion on the pilot study, see the original publications (Appendix 2).

Most importantly, however, the task of silent music reading should be included as standard in eye-movement research in non-textual domains. Relaxing the need to perform allows connecting the findings in this particular task with findings in other domains, while also linking the task of music reading to more general topics of research within cognitive psychology, such as perceptual learning or cognitive styles, to name a few examples.

However, similarly to other openings, also the present experiment yet leaves room for multiple improvements in the research design. First, the viewing time (here 30 s) only allowed rough analyses of the eye-movement processing. Should analyses on the course of the 30-second viewing process have been made, one solution could have been the scan-path analysis often applied in eye-movement studies on perception of static images (e.g. Greene, Liu, & Wolfe, 2012). However, music notation does not present information in a random order, and therefore the viewing process cannot be considered totally random. Thus, the development of analysis tools suitable for examining the course of the viewing processes of music notation in silent reading is a task for future studies. Second, the choice of using an authentic melody, and thus one stimulus only, affected the viewing processes despite the several months between the measurements. It did not seem possible to include control melodies into the explorative design, as controlling for all of the possible factors affecting the viewing processes seemed unlikely to succeed. Thus, the only solution here would be to operate similarly as suggested above in the context of rehearsed reading: one should first search or create a library of suitable stimulus melodies, before administering follow-up studies on the topic. Third, the deliberate choice of leaving the viewing task undefined possibly affected more the novice participants than the amateurs; the former were perhaps more reluctant to describe their intuitive conceptions, being aware of their lack of knowledge on many issues related to music notation. Though the purpose of the experiment was also to explore different silent-reading styles, and this aim could only be reached with such a free-viewing task, future studies could more directly address the effects of different task requirements on visual processing and type of extracted information in the context of music. Fourth, the present experiment instructed the participants to give their descriptions verbally for the purpose of making answering as easy as possible, and allowing the participants to openly hesitate in their responses. Nevertheless, other types of demonstrations of extracted information should also be explored, such as drawings or written responses. In general, exploring the task of silent reading is at its beginning stage, and hopefully, in future, more studies will address this topic.

6. PEDAGOGICAL CONSIDERATIONS AND CHALLENGES FOR FUTURE RESEARCH

This dissertation is among the first presentations of a systematic series of studies on eye movements during different types of music-reading tasks. Through such a line of research, one hopes, after gaining enough cumulative evidence on the roles of the multiple factors influencing the course of visual processing during sight-reading, rehearsed reading and silent reading, to reach a similar understanding on music reading to what has long been a standard in, for example, text-reading research. As similar cognitive mechanisms may underlie the processing of music, text, and other types of stimuli (Palmer, 1997), eye-movement studies on music reading could learn from research conducted in text reading (see Madell & Hébert, 2008). For instance, the role of working memory capacity has not yet been fully explored in the context of music reading, though it is proposed that it plays a role greater than has been thought (Meinz & Hambrick, 2010). In addition, although some techniques from text-reading research have already been tested in the context of music reading (such as the moving-window technique applied in studies on the size of the perceptual span; see Rayner, 1998), the study conditions have not shared the exactness and rigour of the ones already customary in the other fields. Furthermore, on a theoretical level, the cognitive processing of music is quite poorly understood. The only model on the course of saccadic eye movements during a performance task has been presented by Kinsler and Carpenter in 1995. To some surprise, even after nearly two decades this model has not been systematically tested or developed. Therefore models on text reading, such as the E-Z Reader model (Pollatsek, Reichle, & Rayner, 2006), could be modified and explored also in this domain. This dissertation hopefully offers such information on the basic visual processes of music reading that would allow such testing and developing to take place.

Studies on cognitive processes in music reading could help to create suitable learning environments to best enhance learning, or modify the teaching methods presently applied (Gudmundsdottir, 2010a; Lehmann & Gruber, 2006). Already the findings of the present dissertation obviously are contradictory to one method of teaching sight-reading, that is, the teacher's covering of the prior parts of a score while a student is performing, with the intention of "forcing" the student to look far ahead of the current point of performance. In light of the findings of this dissertation, eliminating the possibility of return the gaze towards the currently executed notes may even be detrimental for the performance: even in the simple tasks of Experiment II, the skilled performers first tended to inspect the note they performed during its onset, then moved to inspect upcoming music before returning towards the next performed note (see Figure 1). Thus, it seems that for skilled performers, the natural style of reading while performing is to frequently move back and forth along the score, but the traditional method of teaching is based on the intuitive conception that such eye movements should not occur. More evidence is needed on what exactly lies behind the successful sight-reading strategies, and which strategies truly are the effective ones, but by combining the eye-movement methodology and complementary data, more of these instructional procedures based on intuitive conceptions could be systematically explored.

A second general notion, arising from the experiments of this dissertation, is related to the task of silent reading. The present dissertation wants to emphasize the fact that music notation, despite its main purpose of being a performance aid and a memory tool, does not need to be restricted to, or to be the intellectual property of, only those who are able to perform the symbols. The task of silent reading was administered in the context of a nine-month long music course designed to prepare even the musical novices to teach the subject in elementary school. It was therefore somewhat surprising to note that in terms of the quality of their verbal descriptions, the novices did not benefit from the training any more than did the amateur participants of the experiment. It is clear that during a nine-month long course, it is not possible to gain the performance skills of a fellow student who has had years of practice with an instrument. However, one could assume that even novices in music reading could be taught to "read" music differently, in addition to using notation as a performance aid. It may be that some beginning instrumentalists who struggle at music reading are at a greater risk to cease their hobby (see McPherson, 2005); they are perhaps feeling threatened by the difficulty of reading while performing. A musical score is a complex set of information. Perhaps not enough is done to ease the comprehension of music notation, while teaching emphasizes the successful production of the music and not the intake of information. Exploring these alternative approaches to music reading in more detail could help teachers to understand where the problems lie for students struggling with performing notated music.

After the field of research on music reading has gained sufficient understanding of the visual processing of Western music notation, and identified more of the unique aspects of reading this notation, one hopes, in future, to be able to compare the visual processing of different notational systems. The current system is, as mentioned above, challenging to many learners. Recently, more visually-based systems have been created and tested, and one future task would be to discover what exactly makes these alternative notations easier to comprehend and learn. The multimodality aspect is present in the alternative notations as well, and they also pose the complex task of transforming visual stimuli into corresponding motoric actions. All in all, the systematic series of research on eye movements in music reading is still at a relatively early stage. However, the possibilities offered by current technology and understanding gained of visual processing in other domains suggest that with all good reason, great advances in comprehending the cognitive aspects of music reading, the nature of expertise in this musical task, and the development of educational tools can be attained.

REFERENCES

Adachi, M., Takiuchi, K., & Shoda, H. (2012). Effects of melodic structure and meter on the sight-reading performances of beginners and advanced pianists. In E. Cambouropoulos, C. Tsougras, P. Mavromatis, & K. Pastiadis (Eds.), *Proceedings of the ICMPC – ESCOM 2012 Joint Conference*, Thessaloniki, Greece.

Ahken, S., Comeau, G., Héber, S., & Balasubramaniam, R. (2012). Eye movement patterns during the processing of musical and linguistic syntactic incongruities. *Psychomusicology: Music, Mind & Brain 22*(19). doi: 10.1037/a0026751

Ashby, J., Rayner, K., & Clifton, C. Jr. (2006). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *The Quarterly Journal of Experimental Psychology, 58A*(6), 1065–1086. doi: 10.1080/027249804430000476

Bigand, E., Lalitte, P., Lerdahl, F., Boucheix, J., Gérard, Y., & Pozzo, T. (2010). Looking into the eyes of a conductor performing Lerdahl's «Time after Time». *Musicae Scientiae, Special issue*, 275–294.

Brodsky, W., Henik, A., Rubinstein, B., & Zorman, M. (2003). Auditory imagery from musical notation in expert musicians. *Perception & Psychophysics, 65*(4), 602–612. doi: 10.3758/BF03194586

Brodsky, W., Kessler, Y., Rubinstein, B., Ginsborg, J., & Henik, A. (2008). The mental representation of music notation: Notational audiation. *Perception & Psychophysics, 34*(2), 427–445. doi: 10.1037/0096-1523.34.2.427

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

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*, 55–81.

Clifton, C., Jr., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R. Van Gompel, M. Fisher, W. Murray, and R.L. Hill (Eds.), *Eye Movement Research: A Window on Mind and Brain* (pp. 341–372). Oxford: Elsevier Ltd.

de Groot, A. (1965). *Thought and Choice in Chess.* The Hague: Mouton.

de Groot, A. D., & Gobet, F. (1996). *Perception and Memory in Chess: Studies in the Heuristics of the Professional Eye*. Assen: Van Gorcum.

Drai-Zerbib, V., Baccino, T., & Bigand, E. (2011). Sight-reading expertise: Cross-modality integration investigated using eye-tracking. *Psychology of Music 40*(2), 216–235. doi: 10.1177/0305735610394710

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

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. doi : 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, 266*, 2435–2440.

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

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), 96–123.

Greene, M. R., Liu, T., & Wolfe, J. M. (2012). Reconsidering Yarbus: A failure to predict observers' task from eye movement patterns. *Vision Research, 62*, 1–8. doi*:* 10.1016/j.bbr.2011.03.031

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

Gudmundsdottir, H. R. (2010a). Advances in music reading research. *Music Education Research, 12*(4), 331–338. doi: 10.1080/14613808.2010.504809

Gudmundsdottir, H. R. (2010b). Pitch error analysis of young piano students' music reading performances. *International Journal of Music Education, 28*(61), 61–70. doi: 10.1177/0255761409351342

Hansen, J. P. (1991). The use of eye mark recordings to support verbal retrospection in software testing. *Acta Psychologica, 76*(1), 31–49. doi: 10.1016/0001-6918(91)90052-2

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

Henry, M. L. (2011). The effect of pitch and rhythm difficulty on vocal sight-reading performance. *Journal of Research in Music Education, 59*(1), 72–84. doi: 10.1177/0022429410397199

Hodges, D. (1992). The acquisition of music-reading skills. In R. Colwell (Ed.), *Handbook of Research in Music Teaching and Learning* (pp. 466–471). New York: Schirmer Books.

Hubbard, T. L. (2010). Auditory imagery: Empirical findings. *Psychological Bulletin 136*(2), 302–329. doi: 10.1037/a0018436

Hyönä, J., Lorch, R. F. Jr, & Rinck, M. (2003). Eye movement measures to study global text processing. In J. Hyönä, R. Radach & H. Deubel (Eds.), *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research* (pp. 313–334). Amsterdam: Elsevier Science.

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. doi: 10.1016/j.jecp.2008.04.002

Jacobsen, O. I. (1928). An experimental study of photographing eye-movements in reading music. *Music Supervisors' Journal, 14*, 63 + 65 + 67 + 69.

Jarodzka, H., Scheiter, K., Gerjets, P., & van Gog, T. (2010). In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learning and Instruction, 20*(2), 146–154. doi: 10.1016/j.learninstruc.2009.02.019

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*(4), 329–354. doi: 10.1037/0033-295X.87.4.329

Kaakinen, J. K., & Hyönä, J. (2005). Perspective effects on expository texts on text comprehension: Evidence from think-aloud protocols, eye tracking, and recall. *Discourse Processes, 40*(3), 239–257. doi: 10.1207/s15326950dp4003_4

Kendeou, P., & van den Broek, P. (2005). The effects of readers' misconceptions on comprehension of scientific text. *Journal of Educational Psychology, 97*(2), 235–245. doi: 10.1037/0022-0663.97.2.235

Kendeou, P., & van den Broek, P. (2007). The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts. *Memory & Cognition, 35*(7), 1567–1577.

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

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

Lehmann, A. C., & Ericsson, K. A. (1996). Performance without preparation: Structure and acquisition of expert sight-reading and accompanying performance. *Psychomusicology: Music, Mind & Brain, 15*(1–2), 1–29. doi: 1998-01899-001s

Lehmann, A. C, & Gruber, H. (2006). Music. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge Handbook of Expertise and Expert Performance* (pp. 457–470). New York: Cambridge University Press.

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

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

McConkie, G. W., Zola, D., Grimes, J., Kerr, P. W., Bryant, R. B., & Wolff, P. M. (1991). Children's eye movements during reading. In J. F. Stein (Ed.), *Vision and Visual Dyslexia* (pp. 251–262). London: The Macmillan Press Ltd.

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

McPherson, G. E. (2005). From child to musician: skill development during the beginning stages of learning an instrument. *Psychology of Music, 33*(5), 5–35. doi: 10.1177/0305735605048012

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. doi: 10.1177/0956797610373933

Neisser, U. (1976). *Cognition and Reality. Principles and Implications of Cognitive Psychology.* San Fransisco: W. H. Freeman and Co.

Palmer, C. (1997). Music performance. *Annual Review of Psychology, 48*(1), 115–138. doi: 10.1146/annurev.psych.48.1.115

Penttinen, M., Anto, E., & Mikkilä-Erdmann, M. (2012). Conceptual change, text comprehension, and eye movements during reading. *Research in Science Education.* doi: 10.1007/s11165-012-9313-2 Retrieved from http://link.springer.com/article/10.1007%2Fs11165-012-9313-2?LI=true#page-1

Pike, P. D., & Carter, R. (2010). Employing cognitive chunking techniques to enhance sight-reading performances of undergraduate group-piano students. *International Journal of Music Education, 28*(3), 231–246. doi: 10.1177/025576141037886.

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

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology, 52*(1), 1–56. doi: 10.1016/j.cogpsych.2005.06.001

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

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. doi: 10.1080/17470210902816461

Sloboda, J. (1978). The psychology of music reading. *Psychology of Music, 6*(2), 3–20. doi: 10.1093/acprof:oso/9780198530121.003.0001

Sloboda, J. (1984). Experimental studies of music reading: a review. *Music Perception: An Interdisciplinary Journal, 2*(2), 222–236. doi: 10.1093/acprof:oso/9780198530121.003.0002

Temperley, D. (2001). *The Cognition of Basic Musical Structures*. Cambridge: MIT Press.

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. doi: 10.1080/713756756

van Gog., T., Paas, F., van Merriënboer, J. J. G., & Witte, P. (2005). Uncovering the problem-solving process: Cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology: Applied, 11*(4), 237–244. doi: 10.1037/1076-898X.11.4.237

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

Waters, A., & Underwood, G. (1999). Processing pitch and temporal structures in music reading: Independent or interactive processing mechanisms? *European Journal of Cognitive Psychology, 11*(4), 531–553. doi: 10.1080/095414499382282

Waters, A., Underwood, G., & Findlay, J. M. (1997). Studying expertise in music reading: Use of a pattern-matching paradigm. *Perception & Psychophysics, 59*(4), 477–488. doi: 10.3758/BF03211857

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

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

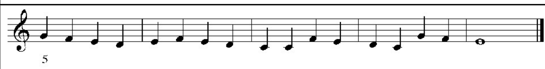
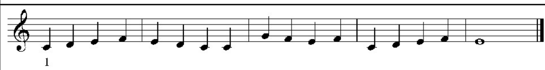
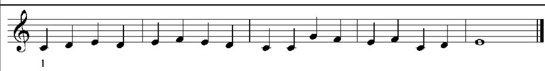
Wristen, B. (2005). Cognition and motor execution in piano sight-reading: A review of literature. *Update: Applications of Research in Music Education, 24*(1), 44–56. doi: 10.1177/87551233050240010106

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

Appendix 1:

Stimulus Melodies of Experiments I, II and III

Melodies in Experiment I composed by Erkki Huovinen. Composer for “Mary Had a Little Lamb” (Experiment II) unknown, variations by Erkki Huovinen, Marjaana Penttinen, Pekka Salonen, and Markku Silander. Composer for “Red Boots” (Experiment III) unknown. All transcriptions by Erkki Huovinen.

Experiment I	1A	
	1B	
	1C	
	1D	
	2A	
	2B	
	2C	
	2D	
	3A	
	3B	
	3C	
	3D	
Experiment II	Original A	
	Variation B	
	Variation C	
Experiment III	“Red Boots”