Measuring Music Reading: A Guide to Assessment Methods

Mary Claire Jensen

A Thesis

Presented in Partial Fulfillment of the requirements for the Degree

Master of Arts in Music

University of Ottawa

Abstract

Music reading is a complex skill. In order to better understand the reading process and evaluate the effects of teaching intervention, it is essential to measure this skill. Research in the field of music pedagogy has provided a number of studies concerned with the measurement of music reading, using varying methods of assessment. However, the corpus of literature is lacking in organization and clarity, in part due to the fact that the assessment methods come from diverse disciplines and the studies themselves may present a number of inconsistencies. Using a research model based on systematic review, the objective of the thesis is to provide an organized synopsis of music reading assessment methods. The thesis has identified and compiled a corpus of 88 relevant studies, with an emphasis on experimental keyboard research in the Western, classical, tonal tradition, though studies with woodwind, brass, percussion, and vocal instrumentation are included as necessary. The assessment methods employed in the studies are classified according to one of three broad categories: test measurements, eye-tracking measurements, and neurological measurements. The purpose of this guide is to be a reference for researchers and educational practitioners, and includes comparison and summary charts and a concluding index.

Key Words: Music Reading, Sight Reading, Measurement, Assessment, Classification

Acknowledgements

I wish to thank many people for their support and encouragement throughout the completion of this project:

To Professor Comeau - your high expectations, standards, and example have pushed me to learn and write more than I ever thought I could. I am thankful for all of your time, honesty, and encouragement.

To Lu, Mikael, and Yixiao - thank you for your friendship, for helping me in so many ways, and for making the Piano Pedagogy Research Laboratory such a lovely place in which to work.

To my thesis committees members Professor Keillor, Professor Moore, and Professor Dineen – thank you for sharing your time and expertise to help me think through and improve this document.

To my colleagues, especially Susan, Sandra, Meganne, Johanna, and Jillian – I am so thankful to have had classes and lunch dates with you. All the very best to each of you in your educational pursuits.

To Ed, Lorraine, Wanda, Mike, and Meganne – thank you for taking the time to read my thesis drafts and to offer feedback.

To all of the teachers who have been examples of excellent teaching and who have developed my love for music and education – Susan Gibbon, Dr. Bill-John Newbrough, Dr. Paul Young, Professor Gilles Comeau – thank you for choosing to dedicate your careers to education.

To Allie, Ottawa and Doylestown friends, and my extended family – thank you for your support in little and big ways. With such a busy two year-stretch, you have helped out more than you'll ever know.

To my very best friend - Greg, thank you for always knowing that I could do this, and for going on this journey with me. I love you.

To my darling baby boy - Henry, you are such a joy and have taught me to slow down and explore. May you continue to love music as you grow. I love you.

To Mom and Dad – thank you for encouraging my love for the piano, and violin, and for all of your sacrifice for my musical pursuits. I am so happy to have been so close to you these past two years and loved having you in our home each week. I love you.

Finally, all glory to God, who has given me all of these opportunities and wonderful people to enjoy them with.

Measuring Music Reading: A Guide to Assessment Methods

Table of Contents

List of Tables

List of Figures

Introduction

The importance of music reading has been widely considered in the literature of the Western, classical tradition, which is the approach emphasized in this thesis. Pedagoges and researchers alike have desired to better understand how one reads and responds to tonal notation. For this reason, there is an expanding body of research with the primary objective to measure music reading. Assessment, through a variety of methods, provides concrete results that elucidate this complex interaction between the brain, eyes, and body. This thesis intends to identify the assessment methods employed in the study of music reading, particularly with keyboard instruments. Using a methodology based on systematic review, which is "an exhaustive overview of the existing literature on a precisely defined subject" (Adèr, Mellenbergh, & Hand, 2008, p. 52), the relevant music reading studies will be classified and presented accordingly.

CHAPTER 1

1. Review of Literature

The review of literature will briefly examine music reading ability, particularly for keyboard musicians. This is followed by an overview of how music reading is assessed. This review will demonstrate that the corpus of experimental music reading studies has received limited attention through researcher reviews, and that there is not an established organizational system for classifying the assessment methods used to study the reading process.

1.1. Music Reading Ability

In the Western, classical tradition, the ability to read notation is considered a valuable and necessary skill for musicians. However, the music reading process is complex, with many factors contributing to its successful execution.

1.1.1. Importance of Music Reading Ability

According to Sloboda (1978a), the ability to read music is an advantage for musicians as he claims that musicians with greater facility for reading have an "immense advantage over other musicians in nearly all walks of musical life" (p. 1). Sloboda (1984) states that the common purpose of music reading is to produce a coherent performance; thus, music reading is an essential tool for the music student to progress and ultimately to perform consistently. Musicians must acquire the skill to read music in order to be able to learn new repertoire without the help of an instructor. According to Henry (2004), "the development of musical literacy would seem essential to the goal of equipping students with independent musicianship skills" (p. 206). For Gilman (2000), it is essential to read music rapidly and efficiently for further career opportunities, such as teaching, accompanying or performing. Salis (1977) writes: "It is no overstatement to say that a person who in some way hopes to earn his living via the piano will be required to sight read 90% of the music he plays, not play it from memory" (p. 1). Lehmann and McArthur (2002) note that although performers no longer sight read in the recital setting, such as Liszt or Czerny did, accompanists and studio musicians possess proficiency in the skill. According to Kopiez, Weihs, Ligges, and Lee (2006), sight reading is one of five basic performance skills.

For keyboard musicians, the skill is increasingly discussed in pedagogical writings. The importance of music reading is featured in many pedagogy textbooks. In *Teaching Piano: A Comprehensive Guide and Reference Book for the Instructor*, Agay (1981) writes, "in piano study the ability to read and play music easily and effortlessly is of prime importance. It is not only an essential factor in the attainment of sound musicianship, but also a potential source of a lifetime of musical enjoyment" (p. 197). The third edition of *How to Teach Piano Successfully* (Bastien, 1995) includes sight reading instruction for first, second, and third year students. *Piano Pedagogy: A Practical Approach* (Parker, 2006) and *The Science and Psychology of Music Performance: Strategies for Teaching and Learning* (Parncutt & McPherson, 2002) include entire chapters on sight reading. There are also chapters in pedagogy textbooks dedicated to reviews of piano method books, with emphasis on the reading portion of the curriculum (Bastien, 1995; Uszler et al., 2000; Baker-Jordan, 2004). Authors of piano pedagogy textbooks have shown a constant preoccupation with the importance of developing music reading skills. Hayward and Eastlund Gromko (2009) attest that music reading skill has great significance and value in music education.

Music reading is also a central component in piano method books. *The Well-Tempered Keyboard Teacher* (Uszler et al., 2000) provides a brief historical overview of music reading and piano method books. During the nineteenth century, the teacher assumed the role of teaching music reading with only the fundamental symbols presented at the beginning of the method books. However, teaching music reading had an increasingly important role in piano method books beginning in the 1920s, and reading strategies have been developed and improved, including three well-known approaches:

middle C, multi-key, and intervallic. A newer approach is the eclectic method, combining two or three of the earlier approaches (Lemay, 2008). Currently in standard piano methods, the development of music reading is the main focus of these books. Recent methods include *Bastien Piano Basics* (Bastien, 1985) which uses the multi-key approach, *The Music Tree* (Clark, Goss, & Holland, 2000), which uses the intervallic approach, *Piano Adventures* (Faber & Faber, 1993), which uses the eclectic approach, *Beanstalk's Basics for Piano* (Finn & Morris, 1998), which uses a varied intervallic approach, the *Hal Leonard Study Piano Library* (Kreader, Wester, Bell, Keveren, & Breth, 1996), which uses a varied intervallic approach, the *Robert Pace Keyboard Approach* (Pace, 1979), which uses the multi-key approach, and *Alfred's Basic Piano Library* (Palmer, Manus, & Vick Lethco, 1995), which uses a combination of the intervallic approach and the middle C approach. A subskill of music reading is referred to as sight reading, generally describing playing at sight without previous study of notation, and it is necessary for musical reading development. Thus, many piano methods also include specific instruction and exercises for sight reading, such as *Bastien Piano Basics* (Bastien, 1985), *Piano Adventures* (Faber & Faber, 1993), *Alfred's Basic Piano Library* (Palmer, Manus, & Vick Lethco, 1995), and there are entire book series dedicated to the development of sight reading skill, including *Musical Encounters* (Bennett & Capp, 1968), *Piano Time Sight Reading* (Hall & Macardle, 1996), *Right@Sight* (Johnson & Evans, 2001), *Four Star Sight Reading and Ear Tests* (Berlin & Markow, 2002), and the *Specimen Sight-Reading Tests* series for the Royal Schools of Music (ABRSM, 2008). According to the preface of the *Four Star Sight Reading and Ear Tests* series, the goal of the sight reading exercises is to develop the sight reading skills of visual learning, tactile sense, aural ability, and analytics.

1.1.2. Complexity of Music Reading Ability

Music reading is complex, with many skills being processed simultaneously. It is a transcription task requiring cognition (Waters et al., 1998) and involves perceptual motor processes (Salis, 1977). This combination of cognitive skills and psychomotor responses (Lee, 2004) can and must be trained. Reading is followed precisely by a mechanical response (Wolf, 1976), which includes “visual, kinesthetic, aural, and aesthetic perceptions and actions” (Salis, 1977, p. 2). Hayward & Eastlund Gromko (2009) state that music reading skill involves both reading and playing, as evidenced by aural discrimination and spatial-temporal reasoning together as one factor, and technical proficiency as another: “Music reading skills draw on auditory processing skills with a spatial component in the formation of an aural image of the sound; playing skills allow the imagined aural image to be made audible” (p. 33). From a psychological viewpoint, music reading also involves memory and problem-solving skills (Lehmann & McArthur, 2002). With regard to reading acquisition for young musicians, McPherson and Gabrielsson (2002) confirm that learning to read is a complex skill; the commencement of which is purposeful, and not instinctive. With music reading and performing, the child must decode the visual stimuli while responding on an instrument.

1.2. Music Reading Assessment

Assessment is essential for evaluating and better understanding the process of music reading. This section first introduces common terms and definitions used for measurement and music reading assessment. It concludes with a historical overview of literature concerned with music assessment.

1.2.1. Definition of Terms

There are many music reading assessments found in research studies, which employ a variety of measurements. A definition of assessment is adopted from Colwell (2002) as a significant collection of data able to provide information for determining or helping to explain cause. Measurements refer to the quantification of an item, characteristic, or behaviour, which are then used for evaluation. Assessments may come from an evaluation or judgment, by expert examiners such as a jury or a music teacher (Elliott, 1982a; 1982b; Eaton, 1978; Lowder, 1973; Lemay, 2008), or through the use of tests or scales that provide measurements or identify and classify errors.

The definitions of music reading and sight reading are not always clear. Hodges (1992) provides a basic definition of music reading: "The process of converting special visual symbols - music notation - into sounds. These sounds may be conceived internally, or they may be produced externally through voices or musical instruments" (p. 466). Another definition includes the quantifier rehearsed preceding the term, and "rehearsed music reading" is presented as the execution of "un- memorized music after sight reading, practicing, and/or imagining the sound of the music from notation" (Miller, 1988, p. 5). However, the term sight reading is more frequently used in the literature, as demonstrated by two standard music dictionaries that do not include a definition for music reading but only for sight reading. *The Oxford Companion to Music* (Latham, 2002) defines sight reading as, "the performance of music from notation that the singer or instrumentalist has not previously seen" (p. 1153). *The New Harvard Dictionary of Music*

(Randel, 2002) defines sight reading as, “the performing of a piece of music on seeing it for the first time” (p. 748). But the term sight reading may have different meanings. Lehmann and McArthur (2002) write, “when musicians speak of sight-reading, not all of them have the same activity in mind. Some might consider only the very first time one reads or plays through an unfamiliar piece to be true sight-reading, while others would allow the definition of sight-reading to encompass play-throughs after more extensive preparation” (p. 135). For example, Miller defines sight reading as an initial execution of notation without previous performance or audiation, which is internal sight reading. But according to Wristen (2005), the situation dictates the type of sight reading task. Pianists may have no preparation or a given amount of time for preparation before performing, but sight reading always means that there is a prohibiting of “total refinement of physical movement in the motor execution phase of the task” (p. 44).

Scientific literature may distinguish between music reading and sight reading, while other research uses both terms interchangeably. Elliott (1982a) provides differentiation between the two terms, with the explanation that sight reading is a specialized form of music reading, but with accurate performance sooner than that of a music reader. Gudmundsdottir (2010b) uses the term music reading with the understanding that it is using a musical instrument to decode the symbols of staff notation. She states that music reading is more comprehensive than sight reading, meaning that sight reading is only referring to music played at first sight, while music reading encompasses all aspects of reading notation. On the other hand, Sloboda (1984) uses the terms interchangeably, but he provides a definition for music reading as the conversion of visual input into a formula for performance. The reader identifies the notes to play in sequence and combination, as well as information about how to perform them. According to Lehmann and McArthur (2002), one can view these terms on a spectrum, with music played at first sight and perhaps at a slower tempo on the sight reading end, while music performed closer to a final tempo and perhaps with rehearsal, would fall at the music reading end. Augmenting the distinction, Wristen (2005) views sight reading tasks as contrasting to music reading tasks of performing repertoire. She considers sight reading an element of learning a piece, with fluid execution only taking place after training and comprehension through continued reading.

For the purposes of this paper, the definition of music reading will be consistent with the definition given by Gudmundsdottir (2010b), which is the decoding and processing of musical notation, whether production requires a motor (i.e. playing on the instrument), verbal (i.e. naming of the notes), or vocal (i.e. sight-singing) response. Sight reading will be defined as a subskill of music reading, and adopts Miller's (1988) definition that there is no previous study or audiation of the musical score. However, the terms music reading and sight reading will be used throughout the paper, respecting the presentation and meaning of the included studies.

1.2.2. Types of Assessment

In music assessment, expert examiners are often used for assessing music reading skills. Here are examples of three studies using expert examiners. Sloboda acted as the sole judge for his 1976b study, and he describes the marking system for sight reading performances:

Altered notes were treated separately. Each note played was scored as an error if it failed to correspond in pitch to the appropriate written note. Rhythm errors, which were slight, were not taken into account, and if a subject corrected himself, his first uncorrected performance was taken as the one for scoring. (p. 234)

Lehmann & Ericsson (1993) provide an example of “a very good and a very poor performance” (p. 189) to two expert judges. These judges are then asked to rate the “overall impression, the technical execution of the performance, and the expressive aspects” (p. 189) on a 6-point scale. The authors note the inter-rater reliability (r = .87), but no other information about them is given. Levy (2001) acted as a judge with two additional judges, to mark miscues from the subjects’ performances. She describes the judges as “educated and experienced instrumental music teachers” (p. 119). Levy also explains: “Because we were able to reliably maintain high agreement (88 - 94%) transcribing performances of different readers and different etudes, I determined that there was no further need for judges to meet together. The rest of the study’s judging was conducted separately” (p. 119). In these examples, the first study has a sole judge, in this case the researcher, using a marking scheme. The second study has two expert judges employing a Likert-type rating scale, and the last study has three judges, one being the researcher, without the use of specified guidelines for analysis. The considerable differences in the evaluations demonstrate the difficulty in standardizing assessment by juries.

Tests or scales for assessments may be developed by field experts and researchers, and vary in their purpose and usage. Three examples of standard published tests are the Watkins-Farnum Performance Scale (Watkins & Farnum, 1954), the Belwin-Mills Singing Achievement Test (Bowles, 1971), and the Musical Aptitude Profile

(Gordon, 1965). These are used frequently, both in research studies and general music education, and incorporate varying measurements. The Watkins-Farnum Performance Scale was designed for instrumentalists and is used primarily in large-group settings with older musicians (Elliott, 1982a; 1982b; Thompson, 1985; McPherson, 1994; Eastlund Gromko, 2004; Hayward & Eastlund Gromko, 2009). This scale measures performance achievement for band instruments, with emphasis on sight reading and general advancement, and with a subsequent scale adapted for string instruments (Farnum, 1969; Comeau, 2009). The Belwin-Mills Singing Achievement Test (Bowles, 1971) has been used in a study with graduate students (Goolsby, 1994a) and is designed to assess sight-singing ability, beginning at an elementary grade 5 level (Roach, 1982). The Musical Aptitude Profile has been used primarily with university students (Erlings, 1977; Micheletti, 1980). It measures tonal imagery, rhythm imagery, and musical sensitivity (Comeau, 2009). While these skills are sometimes studied in relation to reading, this test does not measure music reading. Thus, the Watkins-Farnum Performance Scale is the only standardized test that measures instrumental sight reading. Lemay (2008) adapted this measure for pianists, with appropriate stimuli and a modified grading system and grading chart.

There are also assessment methods for music reading that identify and classify errors. These include the studies of Eaton (1978), Salis (1977), Gilman (2000), and Gudmundsdottir (2003). The rubrics allow for the recording and classification of tonal and rhythmic errors, with special mention of chord errors in Eaton's grading instructions and Salis' error classification. Though measurement scales such as these provide a framework for future research or means for replication, to our knowledge these assessment methods have only been used by the authors themselves in their own research studies. Regarding Eaton's 1978 study, which employs the *Grading Instructions* error quantification rubric, Lemay (2008) writes: "To our knowledge, this method for assessing sight-reading has not since . . . been used either in music reading research or in music education contexts" (p. 21). Gudmundsdottir's *Error Classification* measurement scale was designed for young pianists and has not since been used in research (Lemay, 2008; H. R. Gudmundsdottir, personal communication, February 4, 2015).

For keyboard musicians, the skill of music reading is assessed in many evaluation settings, including piano exams, competitions, and auditions. In exam settings, sight reading tests make up a significant percentage of the total marks. The Royal Conservatory of Music allots 10% for sight reading in the final performance exam score (The Royal Conservatory of Music Official Examination Piano Syllabus, 2015 Edition, The Frederick Harris Music Co., Limited) while the exam board of the Royal Schools of Music (ABRSM) allots 14% for sight reading (http://gb.abrsm.org/en/our-exams/). Sight reading can be the focus of music competitions (Lehmann & McArthur, 2002), and many music festivals include sight reading classes, where participants are scored to read and perform an excerpt or piece, often in front of an audience as well as adjudicators (Kiwanis Music Festival, National Capital Region Syllabus, www.ottawakiwanismusicfestival.com/syllabus.htm). Sight reading can also be a part of auditions (Lehmann and McArthur, 2002) as detailed for acceptance into the piano programs of New York University (http://steinhardt.nyu.edu/music/piano/auditions) and Humber College (http://www.humber.ca/scapa/programs/music-degree/piano-auditions). Thus, music reading is an integral component in the development of musical expertise for students and pedagogues, for amateur musicians, and for professionals.

1.2.3. History of Music Assessment Compilation

Researchers have prepared compilations of music assessment tests for educational organizations or as reference for researchers. An early collection of assessments, experimental studies, and general music publications was assembled by Lehman in 1969, and included published and unpublished tests and measurements spanning from the early 1900s to the mid-1960s. The collection does not contain assessments focusing solely on music reading; however, an assessment of sight-singing measurement (Mosher, 1925) is included, as well as an assessment developed for the rhythm sight reading of instrumentalists (Gutsch, 1964). There is also a test battery to measure the readiness of children for music reading (Hooper, 1966), two studies to determine sight-singing ability (Salisbury & Smith, 1929; Chadwick, 1933), and the presentations of the development and use of a sight reading assessment method for organ music (Stelzer, 1938).

Sloboda (1984) reviewed experimental music reading studies, the purpose of which was to "show that music reading is a genuine species of music perception, by demonstrating the various ways in which musical structure affects performance in readers of varying levels of expertise" (p. 235). The included studies (Bean, 1938; Van Nuys & Weaver, 1943; Weaver, 1943; Sloboda, 1974, 1976a, 1976b, 1977, 1978a, 1983; Salis, 1980; Halpern & Bower, 1979; 1982; Wolf, 1976; Shaffer, 1980; 1981) measure various aspects of music reading. Bean (1938) measured the "span of apprehension", or the number of notes played without error, using a twin tachistoscope mounted on a piano. Subjects would view the stimulus flashed in one of the windows of the device and then perform this stimulus on the piano. Though complete details of the measurement method are not explained, Bean found that experienced performers had a greater span than less-experienced performers. Van Nuys and Weaver (1943) and Weaver (1943) measured eye movements, finding that fixations occur horizontally with contrapuntal music, and vertically with chordal music. Salis (1980) and Sloboda (1976a) measured recall with subjects briefly viewing a musical excerpt and then transcribing it on staff paper. They found that musicians performed better than nonmusicians in a specific time-frame. Halpern and Bower (1979; 1982) measured interference effects, where a subject viewed a visual musical stimulus, participated in a task meant to disrupt recall of the stimulus, and then tried to reproduce the initial display by transcription. They found that musicians did not perform well with visual or auditory interference. Both interference conditions required the subject to classify notes into sets, dependent on their duration or pitch. However, the nonmusician subjects were unaffected by the interference conditions. Shaffer (1980; 1981) measured musical expression and its communication based on the training and experience of performers. The most experienced performers communicated expression most effectively. Sloboda makes use of the data from these studies to determine the differences between good and poor readers and as evidence to frame the assertion that music reading is an input skill of perception.

Hodges (1992) discussed the need for further research in the field, classifying selected studies under the labels of 'basic' research studies and 'applied' research studies, the latter of which were focused on improving music education. This was not a comprehensive review but a summary of important findings.

In another review, quasi-experimental studies of sight reading, including sight reading measurements, were compiled and analyzed by Mishra (2013; 2014). For the sight reading portion, the 2013 paper codes each study by the material used: the Watkins-Farnum Performance Scale, existing repertoire, a published measure for sight reading, or author-composed piece. The compilation also identifies which studies employed a pretest-posttest design. This is the first known compilation of music reading research which emphasizes reading assessment methods, though the intent of the paper is to determine the influence of treatments or interventions for improving sight reading ability and not for the organization and comprehension of reading assessments, per se.

1.3. Research Problem and Objectives

The limitations of music reading research will now be discussed, specifically in the areas of assessment and measurement. Assessment method inconsistencies are described, leading to general limitations for measurement of music reading. This section concludes with the statement of objective for the thesis.

1.3.1. Inconsistency of Assessment

Hodges (1992) states that research has contributed to an awareness of the complex processes involved in music reading, but that further exploration is needed. There is confirmation of this statement with Mishra's 2014 paper, where she writes, "the variability of results from previous research indicates that there is much more to be understood about sight reading" (p. 454). According to McPherson (1994), "there is a continuing need to study the processes involved as instrumentalists sightread music, to compare this method of performance with other aspects of performance, and to substantiate the accuracy of findings with instrumentalists of various abilities" (p. 218).

The formative stage of music reading is an especially important stage of learning, and previous studies have identified the need for research at this phase. There is more research of music reading skill with adult experts than with children; however, reading acquisition is developed during childhood (Gudmundsdottir, 2010b). Tommis and Fazey (1999) identify future research areas that include music literacy with young children, while Wristen (2005) asserts that a comprehensive understanding of the cognitive processes involved in sight reading will be advantageous to the development of pedagogical materials and methods. Penttinen and Huovinen (2011) discuss the need to examine the development of music reading for efficiency and for music educators, specifically with beginning readers: "systematic studies on the steps taken by novice sight-readers learning to perform notated music prima vista are needed in order to offer music teachers more information on which to base their enterprise of promoting students' sightreading skills" (p. 197). There are few music reading studies of young pianists available. What is found in music reading literature is a study of child pianists by Drake and Palmer (2000), where all subjects learned a piece of music in eleven trials, with the purpose of following cognitive changes with short- and long-term learning. Errors were measured by computer, and encompassed errors of pitch, duration, or a combination of pitch and duration. Corrections and pauses were also considered errors, and scored appropriately. Drake and Palmer found that the following variables improved through practice, or short-term learning: speed, accuracy, timing and continuity of rhythm, and the ability to plan ahead. With increased skill, or long-term learning, they found that the range of anticipation increased, resulting in fewer disruptions of rhythm. Tommis and Fazey (1999) studied preschool-aged children with the research goal of comparing two teaching methods for the introduction of conventional Western notation. They measured the children's understanding of notation through a pretest and posttests using a variety of age-appropriate apparatus: author-designed paper keyboards, an electronic keyboard, note flash cards, and short melodies for the children to ‘perform’ by pointing or playing. This study is one of the small number of examinations of very young children and music reading, with the stimuli using only five keyboard notes in the treble clef; however, Tommis and Fazey found that both methods were successful and the subjects acquired music literacy skills over the 8-week instructional program. Gudmundsdottir (2010a) studied pianists aged 6 through 13 as they played simple pieces. She measured errors of pitch, including erroneous, redundant, and omitted pitches. Gudmundsdottir found that the younger pianists made more incorrect and redundant pitch errors, and that all subjects made more left hand errors. She also found that regardless of formal instruction, music literacy skills continue to develop through childhood, as evidenced by the older subjects making fewer errors. The data from these studies needs validation through continued, replicated, and expanded study. According to Mishra (2014), sight reading research is a growing field, but without organization: “replication studies are exceedingly rare in the field, and the sheer number of variables makes it impossible to test (or retest) all possible variables within one study. This has led to widely disparate results with little to connect them” (p. 453). In order to do research, consistent measurement methods for assessment are essential.

1.3.2. Limits of Music Reading Assessment

Lemay (2008) examined three assessment methods and found that they varied greatly in their results. The purpose of her study was to better understand the assessment methods available for sight reading performance, and to achieve this Lemay compared three of the methods with their assessment of eight individual performances. The methods to be compared consisted of the Watkins-Farnum Performance Scale (Watkins & Farnum, 1954), Scoring Algorithm (Gilman, 2000) and evaluations by expert examiners. Lemay writes, "these methods have . . . never been compared to determine if they provide analogous assessments of reading accuracy in performance. When assessment methods used in research vary considerably, it is difficult to provide reliable and analogous data" (p. 31). Each method evaluated each performance, and the subjects were ranked, according to that particular method. The assessments were tailored to young pianists with appropriate stimuli, and the subjects for the comparison study were pianists between the ages of 11 and 16. The subjects performed increasingly difficult sight reading exercises, eleven in total, or until they could no longer play without error in each measure. Each system of measurement varied in its degree of severity or leniency, and thus the subjects' ranking varied for each method. The assessments did not value the same components of sight reading though all were designed to assess sight reading skill. For example, articulation and expression errors were only considered by the expert examiners and not by the standardized evaluation methods. This is an example of confusion with three assessment methods in a controlled study with broad implications: validity cannot be established for use in educational and research settings, music reading studies which have already used these methods to rank subjects in terms of their reading ability may vary if replicated, and, as shown in this study, the methods can not provide analogous measurements for continued assessment in sight-reading performance (Lemay, 2008).

Additional inconsistencies between one music reading study to another must also be considered. Studies are comprised of varying subjects, stimuli, and instrumentation. With keyboardists, several studies use only expert pianists (Waters, Townsend, & Underwood, 1998; Salis, 1980; Kopiez & Galley, 2002) or provide few descriptors to qualify their participants, such as Sloboda's 1974 study, which refers to subjects only as music students. Studies use wide-ranging terms to describe the pianists' reading skill: Sloboda (1984) uses the terms good and poor, Waters et al. (1998) uses the terms more-skilled and less-skilled, Kostka (2000) has subjects in different levels of piano study, referred to as Level I and Level III. Stimuli may be author-composed repertoire, published repertoire, or use of a standard measurement method (Mishra, 2013). Stimuli for piano reading studies may be for only one hand (Waters & Underwood, 1999; Stewart, Walsh, & Frith, 2004 [experiment 1]; Sloboda, 1974; Penttinen & Huovinen, 2011), or two hands (Lehmann & Ericsson, 1993, 1996; Gudmundsdottir, 2010a). A study of younger children in the early elementary grades did not use traditional notation in its stimuli, but instead an author-created notation that appeared similar to traditional notation but did not aid in the learning of treble and bass staves (Pick, Unze, Metz, & Richardson, 1982). Instrumentation can include the use of a keyboard-instrument (Salis, 1980; Kopiez & Galley, 2002), wind instruments (McPherson, 1995; Goolsby, 1994a; MacKnight, 1975; Palmer, 1976; Bebeau, 1982), or without the use of an instrument (Schön & Besson, 2002; Stewart, Walsh, & Frith, 2004 [experiment 2]). These inconsistencies and lack of information demonstrate how difficult it can be to replicate a study, compare studies, or to validate assessment methods. Historically and currently, music measurement is a field in need of organization and stability.

At the time of his study, Lehman (1969) discusses the many drawbacks of music evaluation: the varieties of tests show the complexity of evaluating music skills and the weaknesses of previous tests. Lehman stresses that tests are often inadequate, with few solutions to prepare evaluations to attain the appropriate results. His wish was that the compilation of music assessments and studies would provide direction for future research. This comment is still valid today, as research is fruitless when studies do not use the most suitable measurements available. This appears to be particularly true of music reading assessment research, as researchers are often unaware of the measurement techniques available. As seen with Sloboda (1984), Hodges (1992), and Mishra (2013; 2014), compiling and analyzing research studies can offer organization to a large body of work, provide valuable data, identify weaknesses or limitations, and guide future research. However, it appears that there has not been an attempt to undergo a comprehensive and methodical review of all available research concerned with music reading. Our research will attempt to provide organization to music reading measurement, with a special emphasis on keyboard performance. The aim is to help to clarify and standardize reading measurements in order to obtain valid and reliable results in future research.

1.3.3. Statement of Objective

The previous review of literature and the described limitations of music reading assessment and its lack of organized presentation help to establish the objective of this study: to standardize the existing measurement methods used to assess music reading. This will be achieved by:

1. Conducting a search on the current body of literature based on a systematic review model,

2. Identifying the different measurement methods for the assessment of music reading,

3. Developing a classification system for the methods,

4. Providing a similar and consistent body of information on each method, and

4. Providing an overview of the different research studies that have used such measurement methods.

The descriptions of reading assessment methods are widely scattered among the fields of social science, neurology, music education, and psychology. Music researchers may be unaware of the wide variety of measurement practices available, which is unfortunate as it results in less scientific data collection due to research employing different methods or not using the best available method. This thesis was written with the intention of compiling the assessment means. We reviewed the leading measurement methods and we now provide a detailed presentation of each. Therefore, this thesis is a guide that reviews all the different measurement systems used to assess music reading.

CHAPTER 2

2. Methodology

The focus of our project is music reading assessment, and the objective is to identify and classify music reading measurement methods. We will undertake a review of existing studies based on a systematic review approach. According to Nunn and Hill (2016), a systematic review comprises:

1. Establishing a research problem
2. Conducting a comprehensive search for appropriate and relevant information
3. Examining the information for specific details
4. Considering the information within a given context to evaluate its quality
5. Examining and comparing the body of research collected, to give an overall result

For this study, the research problem is the need for a framework that will provide stability and consistency for subsequent classification. The classification of music reading measurements should then provide organization for existing and ongoing music reading research. Upon completion of the systematic review, we will identify the measurement methods and construct a framework based on the methods, with each study categorized and summarized within the established framework.

2.1. Establishing a Research Problem

The research problem for this study was identified after examining the review and compilation literature available. It is evident that there is not a classification system in place, nor any previous attempt to create a system which incorporates the body of experimental studies concerned with music reading assessment methods.

2.2. Conducting a Comprehensive Search

A comprehensive search of studies concerning music reading was completed. This included published studies in journals, music texts, theses, and dissertations.

Initially, an existing reference list of studies on music reading was used, as provided by the Piano Pedagogy Research Laboratory at the University of Ottawa. To obtain additional references, a key word search was completed using the following databases: the University of Ottawa's Classic Catalogue, the University of Ottawa's Music Index database, Scholar's Portal, ERIC, JSTOR, RILM, Proquest Thesis and Dissertation, International Index to Music Periodicals, Google Scholar, and PsycINFO. The search terms included the following, individually or in combination: music reading, sight reading, piano reading, reading assessment, testing, reading, measuring, assessment, evaluation, study, cognition, brain, children, pedagogy, notation, score, piano, novice, expert. The references obtained from this search were explored and analyzed to identify the studies that were relevant, based on the clarity of the measurement methods the studies were presenting. The conceptual or methodological value and quality were also of consideration. Studies were not retained if they were irrelevant to music reading or to measurement, or if they were dated before 1960, with the exception of any studies that were of historical significance in the field.

Initially, 103 studies were identified as providing putative assessment methods for music reading and thus appropriate for our research, ranging from the years 1966 to 2015. However, after further examination, a number of the preliminary studies were found to emphasize other areas of music comprehension, such as perception, performance, or the relationship between music and language literacy. These studies were consequently removed as they were determined to be too far from the focus of the thesis. In addition, twenty-one relevant studies were also found throughout the period of research, and were added to the collection. The final number of classified studies is 88. This includes published papers (n = 70, 80%), doctoral dissertations (n = 15, 17%), published books (n = 1, 1%), and published symposium presentations (n = 2, 2%).

2.3. Examining the Information

Particular information will be extracted and examined from each study, as the information is available. As well, each study has a corresponding chart with the same information in Appendix B. The charts contain a condensed description of the study, divided by labeled headings, and have been written with a combination of the author's words and text directly copied from the original study, in order to remain true to the context of the publication but concise for reference purposes. This selected information is indicated in Table 1.

Table 1. *Framework for chart presentation.*

Title
Author
Year
Purpose
Measurements
Assessments Used
Description
Subjects
Apparatus Used
Stimuli
Apparatus
Method
Results
Reliability
Validity
Reference
See Also

We then proceed with a synthesis and comparison of information. Each study is introduced in the thesis body by a brief table that outlines relevant information for the purposes of comparison and understanding.

Table 2. *Example of summary table framework for test measurements.*

Test Measurements Study A, Year	
Type	
Test Construction	
Production Method	
Participants	
Stimuli	
Pre- and Post-Tests	
Control Group	
Scoring System	
Duration of Instruction	(optional category)

Table 3. *Example of summary table framework for eye-tracking measurements.*

Eye-Tracking Measurements Study A, Year	
Apparatus	
Production Method	
Participants	
Stimuli	
Pre- and Post-Tests	
Control Group	
Measurement	

Table 4. *Example of summary table framework for neurological measurements.*

Neurological Measurements Study A, Year	
Apparatus	
Production Method	
Participants	
Stimuli	
Control Group	
Measurement	

The chapter concludes with a synthesis of the tables.

Table 5. *Example of summary tables synthesized for eye-tracking measurements.*

	Apparatus	Production Method	Participants	Measurement
Eye-Tracking Study A, Year				
Eye-Tracking Study B, Year				
Eye-Tracking Study C, Year				

2.4. Considering the Information within a Given Context

The selection process for the established classification system, including the chosen subcategories, is described below. It includes a framework for the presentation of each measurement method in its entirety, a framework for the summary and synthesis tables throughout each chapter, and a framework for the summary charts located in Appendix B.

2.4.1. Selection of the Broad Categories of the Framework

To establish a classification framework, *The APA Handbook of Testing and Assessment in Psychology, Vol. 1: Test Theory and Testing and Assessment in Industrial and Organizational Psychology* (Geisinger, 2013) was used as a resource. In this compendium, tests are categorized in a variety of ways: by their administrative purposes, by the conditions under which the test is being administered, or by the information gathered. Another resource was consulted from the field of music. Hodges' (1992) review of sight reading research categorized studies under the headings of eye-movement, the relationship between music reading skill and related variables, the pedagogy of music reading, and error-detection in score reading.

Though these resources were valuable for understanding of how to commence the classification of a large body of work, after examining and summarizing the 88 music reading studies into separate charts, it was clear that there were three large categories in which to classify the music reading assessment methods. Music reading skill or process is primarily measured with the use of tests and scales, by studying eye-movement and response, or by neurological observation. Thus, the broad categories of the framework are test measurements, eye-tracking measurements, and neurological measurements.

2.4.2. Selection of the Subcategories of the Framework

For the subcategories of test measurements, we attempted a number of organizational systems. A large number of studies employed The Watkins-Farnum Performance Scale (Watkins & Farnum, 1954), as this is a published and standardized test for the evaluation of instrumental music reading. From this, we surveyed additional published music tests, which are primarily tests of aptitude or achievement without a set reading component. The study of aptitude and achievement tests led us back to our collected studies, where we noticed how many studies employed a series of published tests in order to form a conclusion, and how often a researcher constructed a test for the purposes of their inquiry because a specific music reading measure was lacking. Thus, the final subcategories include general music tests of aptitude or achievement, and music reading tests that are either standardized and published, or author-composed. The available studies in test measurements can be classified in one of four areas: by their focus on cognitive process, on instructional methods and effects, on correlation with music reading, or on validation of their proposed assessment method.

For eye-tracking measurements, it was evident that the studies could be subdivided by their use of apparatus: by viewing devices such as the tachistoscope or tachistoscopic-like techniques, and eye-tracking devices. Examination of eye-tracking devices revealed that a large number of eye-trackers are based on infrared technology. Thus, the final subcategories are measurement by viewing devices, by eye-tracking devices (infrared-based and other), and by electrooculography. Again, the studies in this field can be further classified by their research focus, though the majority appear to be concerned with cognitive process.

For neurological measurements, it was also clear that subdivisions by neurological apparatus were most suitable. Neurological measurements can be obtained by the use of positron emission tomography (PET), electroencephalography (EEG), magnetoencephalography (MEG), magnetic resonance imaging (MRI), and functional magnetic resonance imaging (fMRI) (Florh & Hodges, 2002). As well, neurological assessments diagnose brain activity in response to interference or physical impairment. Neurological studies can be further classified, with the overwhelming majority of studies focusing on cognitive process.

Table 6. *Framework with broad category headings and subcategory headings.*

Test Measurements
General Music Tests
Tests of Aptitude
Tests of Achievement
Music Reading Tests
Sight Reading or Rehearsed Reading
Standard Test
Author-Composed Tests
Sight-Singing
Standard Test
Author-Composed Tests
Music Reading Studies Using Tests
Cognitive Process Studies
Instructional Studies
Correlational Studies
Test Validation Studies
Eye-Tracking Measurements
Viewing Devices
The Tachistoscope
The Light-Out Technique
Eye-Tracking Apparatus
Various Eye-Trackers
Infrared-Based Eye-Trackers
Electrooculography
Music Reading Studies Using Eye Movement Methods
Neurological Measurements
Positron Emission Tomography (PET)
Electroencephalography (EEG)
Magnetoencephalography (MEG)
Magnetic Resonance Imaging (MRI)
Functional Magnetic Resonance Imaging (fMRI)
Neurological Assessment of Injury and Music Reading
Music Reading Studies Using Neuroimaging Methods

2.4.3. Framework for each Chapter

The criterion for the overall chapter framework is established from a current compilation of health assessments. In *Measuring Health: A Guide to Rating Scales and Questionnaires* (McDowell, 2006), quality, published, and inclusive measurements are presented with information on established reliability and validity. According to McDowell (2006), “The aim has been to provide enough information to permit the reader an informed choice among instruments . . . the goal here is to provide a representative picture of studies that have evaluated each measure, and consequently our reviews vary in length" (p. 5). McDowell's references include an historical overview of the measurement techniques to better understand the evolution of the methods followed by a comparison of the different measures, with the intention of helping the reader choose the most appropriate test for their needs. The simple measurements are described first, followed by increasing complex measurements, which generally follows in a chronological order. Each review varies in length due to the complexity of the measurement and the available information. Each chapter concludes with a summary, direction for further study, and additional measurements that deserve mention. Based on the framework in *Measuring Health*, we will include a historical overview at the beginning of each chapter in the thesis body, followed by descriptions of the various measurement apparatus, and then describe each study in chronological order, when possible:

Table 7. *Framework with history, apparatus descriptions, and study descriptions.*

Eye-Tracking Measurements
History
Description of Apparatus 1
Description of Apparatus 2
Music Reading Studies Using Eye Movement Methods
Apparatus 1
Study A
Study B
Apparatus 2
Study A
Neurological Measurements
History
Description of Apparatus 1
Etc.

We have established the following for the measurement methods: 48 studies employ test measurements (55%), 25 studies employ eye-tracking measurements (28%), and 15 studies of neurological measurements (17%). Note that the studies of Salis (1977), Thompson (1985), Goolsby (1994a), and Gilman (2000) are classified under both test measurements and eye-tracking measurements categories, but are counted only once for the above summation, according to their primary research. Thus, the studies of Salis and Thompson are counted for test measurements, while the studies of Goolsby and Gilman are counted for eye-tracking measurements.

The classification process will help to establish areas of music reading assessment that are lacking. According to Hubley and Zumbo (2013), "yet the various choices made throughout the measurement and assessment framework have important consequences for the definition, quantification, and use of tests and measures and the decisions that are based thereon" (p. 18). Each measurement method will be presented within the framework in the following manner:

Table 8. *Example of complete chapter framework.*

- History
    - Description of Apparatus 1
    - Description of Apparatus 2
- Music Reading Studies
    - Apparatus 1
        - Summary Table
        - Study A
        - Summary Table
        - Study B
    - Apparatus 2
        - Summary Table
        - Study A
        - Summary Table
        - Study B
- Summary
- Synthesis of Summary Tables

2.5. Examining and Comparing the Body of Research to Give an Overall Result

The conclusion of the thesis will reflect on the results of the review and classification process, and the results from comparison within each assessment method and between the assessment methods. Appendix A presents the classified studies by category, with full and related references. The charts are ordered alphabetically in appendix B. Finally; a thorough index allows the reader to locate a specific measurement, as well as key words or names.

CHAPTER 3

3. Test Measurements

Measurement and assessment through testing further the progression of psychology research, under which music is studied (Geisinger, 2013). Particularly for music, testing provides direction for curriculum development and focus for training purposes (Gordon, 1971; Erlings, 1977; Shuter-Dyson & Gabriel, 1981), further understanding of relationships or correlations between music reading ability and musical aptitude or skills, and the collected data can be used to create predictive models (Erlings, 1977; Miller, 1988).

3.1. Test Descriptions

Tests will be described with their historical significance, if applicable, and by the nature of test. Descriptions of general music tests include tests of aptitude, in regards to how they relate to or employ music reading, and tests of achievement, which include areas of notational measurement but not played performance. All of the described general music tests are published and have been used at one point, or currently, in educational settings. Music reading tests include standard or author-composed tests, and measure cognitive process, instructional effects, or subskills of music reading that correlate with reading achievement. Music reading tests may also be author-constructed and presented in a study for validation purposes.

All tests are constructed with a scoring apparatus. Results are obtained by expert evaluators or by scoring technology. Expert evaluators include the researchers who conduct the experimentation for validation or for research. Although it is rare, the researchers may administer and score the tests unaided, such as with Sloboda (1976b) or the modified WFPS scoring by Miller (1988). More often, researchers employ research assistants, doctoral or graduate students, or experienced music teachers to assist in scoring. There are times when the researcher cannot be present during test scoring in order to maintain the integrity of the research. Expert evaluation will naturally have a degree of subjectivity, though detailed scoring systems and reliability procedures encourage consistency.

Other scoring systems primarily involve computer programs that collect, record, analyze, and print data. Henry (2015) writes, "The landscape of educational assessment is continually changing. Advancements, particularly in the mode of delivery, have occurred at an accelerated pace as computer access in the schools has also increased" (P. 58). With regard to music assessment, published and routine instrumental testing methods, with live or recorded performances, have been the forerunner for vocal testing methods, such as the WFPS (Watkins & Farnum, 1954; Farnum, 1969). This is also true of recent assessment methods using technology; in particular, music software. Henry notes that such software can evaluate performances, including sight reading performance, and is continuing to be developed for vocal sight reading testing. Potential advantages of using technology-based assessment methods include the fact that it can be used individually and possibly without a test administrator, that it can impose an exact tempo, and that it can record the performances and provide feedback. Zhukov (2014a) used a software program to score and analyze piano sight reading, and states that the scoring algorithm within the program "provided greater consistency and reliability of analysis of pre- and post-tests than manual marking of errors utilised in many previous sight-reading studies" (p. 76). Zhukov specifically names the studies of Banton (1995), Betts and Cassidy (2000), Kostka (2000), Meinz and Hambrick (2010) and Penttinen and Huovinen (2011). Computer software has been employed to score piano music reading tests, including Lehmann & Ericsson (1993; 1996), Drake & Palmer (2000), Kopiez, Weihs, Ligges, and Lee (2006), Kopiez and Lee (2006; 2008), Zhukov (2014a), and an adaptation of a sight reading test by Stewart, Walsh, and Frith (2004).

For music reading, there are numerous production methods. Performance on an instrument or sight-singing are most common when measuring music reading ability. Verbal response and transcription may be used to measure music reading ability. Halpern and Bower (1982) also employed transcription, which involves the subject viewing a stimulus for a brief period and then notating it from memory. For rhythm-reading, subjects may clap, give a verbal response, or use a percussive instrument. Finally, the method of repetition, or many reiterations through performance, has been used in two studies (Pick et al., 1982; Shehan, 1987). In both cases, the measurement method consisted of counting the number of performance repetitions needed to achieve a perfect execution of the stimulus.

3.1.1. General Music Tests

According to Fourie (2004), the study of how the brain processes music has been undertaken for well over a century and from many disciplines, including the field of psychological testing. Psychological testing has its origins in China in 2220 B.C. (Treichler, 2013; Gregory, 2007). However, the precursor to present-day testing developed in the late 1800s in Europe, by researchers such as Sir Francis Galton and Wilhelm Wundt. There is published evidence of mental testing by the American James McKeen Cattell in 1890 (Gregory, 2007). With regard to musical talent, Galton studied how musical talent was inherited, and psychologist Carl Stumpf conducted early research using tasks such as judging pitch or singing a played pitch. However, Galton was limited by the absence of controlled testing methods, and Stumpf's work was not standardized nor published (Shuter-Dyson & Gabriel, 1981; Treichler, 2013). In America, W. E. Scripture established a psychological laboratory at Yale University, modeled after the European laboratory of Wundt. Here, he studied the senses, including vision and hearing (Shuter-Dyson & Gabriel, 1981). Also in America, Carl E. Seashore and Raleigh Drake helped to expand the field in the early 1900s, and were particularly concerned with the prediction of musical ability (Treichler, 2013). Throughout the 1930s and 1940s, the Oregon Music Discrimination Test by Kate Hevner and Drake's Musical Aptitude Tests were developed as effective musical test batteries. At this point, tests of performance achievement were also designed and tested, such as Watkins (1942) study, as described below. According to Shuter-Dyson and Gabriel (1981), since the 1940s, studies have validated many of these tests and thus the tests have been used in education and for research, with data analysis provided by improved statistical methods. The authors also write, "new tests, especially of aural achievement in music and of the musical ability of children...have been devised...so important have become the areas of perception, memory and cognition of music and its physiological correlates (p. xiii).

Shuter-Dyson and Gabriel (1981) note that there are two broad categories for music evaluations: aptitude tests, which are predictive of skill, and achievement tests, which comprise learning assessment. The following tests are measures of musical aptitude or achievement without a traditional performance component. These evaluative tools encompass the vast majority of available published tests and are useful for research purposes: the use of an existing, published, and standardized test allows for gathering continued data and provides a foundation for novel tests. The tests are included because they are part of a series of tests with the objective to better understand music literacy, or to correlate subcategories of music comprehension or skill with general reading ability. As such, specific sections of the following tests that pertain to music reading, or that measure music reading, are explained in greater detail.

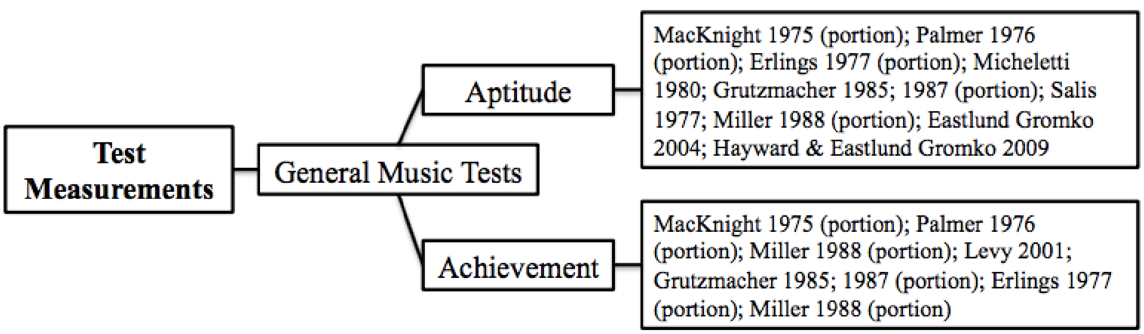


Figure 1. *General music tests, subcategories, and studies.*

3.1.1.1. Tests of Aptitude

Gordon (1971) states that the purposes of aptitude tests are to assess complete aptitude, to identify strengths and limitations of an individual, to observe development, and to compare an individual with their peers. Since musical achievement tests "have aural items requiring musical aptitude as well as a knowledge of notation" (Shuter-Dyson & Gabriel, 1981, p. 38), musical aptitude is an important part of music reading.

The Drake Musical Aptitude Tests

Raleigh Drake was a musician and psychologist, devising the first of the Drake Musical Aptitude Tests in 1933, and the second in 1954. The first test concerns musical memory and interval discrimination. The second is an assessment of rhythm (Shuter-Dyson & Gabriel, 1981). There are two forms of each. According to Shuter-Dyson and Gabriel, the published tests are frequently used, due to the fact that the musical memory test is admired and supported, and the rhythm test "is the only...test available that specifically measure[s] ability to keep in time" (p. 21). In music reading literature, Miller (1988) has used the rhythm test, while Salis (1977) has used both tests. Neither of the tests requires music reading; instead, both have listening components for a same-different choice or counting in silence, respectively (Salis, 1977; Shuter-Dyson & Gabriel, 1981; Miller, 1988).

The Musical Aptitude Profile (MAP)

Created by Edwin Gordon and published in 1965, the Musical Aptitude Profile (MAP) is a test battery, described by Gordon as "an eclectic battery, drawing from both gestalt and atomistic theories" (p. 19). It is administered over three days for students in a group setting (Shuter-Dyson & Gabriel, 1981). The MAP is based on the concept that music aptitude is inherent, and therefore musical training cannot influence the resulting score (Gordon, 1986). It has three subtests: tonal imagery, which includes melody and harmony, rhythm imagery, which includes tempo and meter, and musical sensitivity, which includes phrasing, balance, and style (Shuter-Dyson & Gabriel, 1981). McKnight (1975), Palmer (1976), Erlings (1977), Micheletti (1980), and Grutzmacher (1985; 1987) have used the entire Musical Aptitude Profile, or portions of the MAP, in their respective studies. None of the subtests require music reading; instead, musical recordings are presented to the subjects who respond on an answer sheet. However, following a three-year longitudinal study of 250 students administered the MAP, “The profile was shown to be a valuable predictor of achievement” (Shuter-Dyson & Gabriel, 1981, p. 27). For Erlings (1977), she concluded that portions of the MAP correlated significantly with “keyboard sight reading achievement” (p. 16).

The Advanced Measures of Music Audiation (AMMA)

The Advanced Measures of Music Audiation (AMMA) was published in 1989 by Edwin Gordon, and contains test items similar, though with increased difficulty, to those found in the 1965 Musical Aptitude Profile by Gordon (Comeau, 2009). The Primary Measures of Music Audiation and the Intermediate Measures of Music Audiation precede the AMMA, and all three tests measure the source and stability of music aptitude (Gordon, 1986; Comeau, 2009). There is no music reading component to the battery, specifically. All items are presented by a musical recording, and the subjects respond after listening (Gordon, 1986). Eastlund Gromko (2004) used the AMMA to measure tonal and rhythmic audiation, because she had found that sight reading was associated with audiation, among other factors. Hayward and Eastlund Gromko (2009) used the AMMA to measure “aural discrimination of patterns” (p. 31), and subjects were required to determine the similarities or differences between two musical patterns, and, if a difference was found, which element (tonal or rhythmic) had been modified.

3.1.1.2. Tests of Achievement

Regarding achievement testing, Shuter-Dyson and Gabriel write, “attainment tests may take the form of questionnaires on musical knowledge, or of scales against which vocal or instrumental performance can be compared. Most prognostic tests so far developed deal only with the aural side...some tests, though mainly aural, require a knowledge of notation” (p. 12-13).

The Aliferis Music Achievement Test (MAT)

The Aliferis Music Achievement Test (MAT) is a published and standardized test designed for music students entering post-secondary music studies, with the first level introduced in 1954, followed by a second level in 1962 (Aliferis, 1957; Shuter-Dyson & Gabriel, 1981). According to Aliferis (1957), the test measures the “ability to hear with the inner ear what is seen in notation and to visualize the notation of music that is heard” (p. 6), or *auditory-visual discrimination*. All stimuli are presented on the piano, and the first level test has melody, harmony, and rhythm subtests. The second level has subtests of harmonic elements, melodic idioms, and rhythmic idioms. Melodic elements are represented by intervals, and rhythmic elements equal the duration of one beat. The idioms consist of four-note patterns (Aliferis, 1957; Shuter-Dyson & Gabriel, 1981). According to Aliferis (1957), this test correlates with the ability to sight-sing and notate dictation. The Aliferis MAT has been used by Erlings (1977) where subjects also completed two piano sight reading subtests, after which the MAT was scored. One of Erlings’ conclusions was that the Aliferis MAT correlated notably with sight reading achievement.

The Bentley Measures of Musical Abilities

The Measures of Musical Abilities is a test battery constructed by Arnold Bentley and published in 1966 (Rowntree, 1970; Shuter-Dyson & Gabriel, 1981; Comeau, 2009). It is best suited for children and the results help to understand children's processing of musical concepts and responses (Rowntree, 1970). Law and Zentner (2012) clarify that the measure is designed for ages 9 to 11, specifically, and add that it is a 'convergent' and 'criterion' test, but not a 'predictive' test (p. 1). The battery measures basic music perception and discrimination that is either innate or developed without formal training (Musumeci, 1997; Comeau, 2009). There are four subtests, which require written response (Musumeci, 1997), including pitch discrimination, tonal and rhythmic memory, and chord analysis (Rowntree, 1970; Comeau, 2009). There are no tasks involving music reading; however, Rowntree (1970) suggests significant correlation of The Measures of Musical Abilities with instrumental and singing ability. Miller used this test battery in his 1988 study. In the measure of musical memory, subjects determined the similarities or differences between a target pitch or rhythm phrase and a second phrase. In the measure of music aptitude, subjects compared the pitches of two notes, and in a second task, determined the number of notes in a played chord. Generally, this test battery is no longer used as it is considered insufficient in its reliability and inconsistent in measurement objectives (Rowntree, 1970; Comeau, 2009; Law & Zentner, 2012).

The Colwell Musical Achievement Test (MAT)

The Musical Achievement Test (MAT) was published by Richard Colwell in 1969 and 1970. The tests measure auditory elements of musicianship for students from grades 3 through 12, with the purpose of determining which students will most benefit from instrument training, as well as for appraising and improving music curriculum (Shuter-Dyson & Gabriel, 1981). According to Kornicke (1992), this is one of the few published tests that measures “aural imaging” (p. 154), which is important for ear training. Colwell (1970) writes, “MAT is an aural test, for music is aural as an art, skill, and activity...music reading is greatly aided by an aural concept of scale; it seems unlikely that students will learn to hear, recognize, or sing intervals without the ability to recognize scalewise patterns” (p. 62).

There are four tests in the MAT, which can be administered collectively or individually (Shuter-Dyson & Gabriel, 1981). Of the music reading literature, all researchers have primarily employed Test 2 of the MAT in combination with tests by other authors (MacKnight, 1975; Palmer, 1976; Miller, 1988; Levy, 2001). Test 2 involves three subtests (Shuter-Dyson & Gabriel, 1981). The first subtest of Test 2, *major-minor discrimination,* requires subjects to identify major or minor chords and phrases performed on piano. The subtest of *feeling for tonal centre* requires subjects listen to a cadence or phrase and name the central note by choosing one of three pitches played afterwards. The subtest of *auditory-visual discrimination* requires music reading, as subjects read notation while listening to the pitches and rhythm, in order to identify discrepancies between the performance and the score. Though not used in the literature reviewed by this thesis, it is of note that Test 3 of the MAT has a *pitch recognition* subtest, which Colwell identifies as directed towards a “single facet of music reading, and that is correct interpretation of intervals” (p. 68). In this subtest, the first note of an interval is played while the subject reads the notation. The second note of the interval is then written on the score, and the subject chooses the correct note from three performed tones. As well, Test 4 of the MAT has an auditory-visual discrimination subtest. However, Colwell (1970) remarks that though these specific subtests are suggestive of music reading ability, “it is really more elementary than genuine music reading, and more closely correlates with the ability to follow a line score. The rhythm and pitch sections are separated, so that the test furnishes diagnostic information” (p. 65). For the complete Colwell MAT, there is high reliability with the complete score of all tests (Shuter-Dyson & Gabriel, 1981).

The Iowa Tests of Musical Literacy (ITML)

Edwin Gordon (1970) wrote the Iowa Tests of Musical Literacy (ITML). Described as a “six-level achievement battery” (Sidnell, 1973, p. 54), it is for group administration, with the presentation of recorded stimuli (Grutzmacher, 1985; Schleuter, 1974). Gordon (1970) describes the purpose of the tests: “to sequentially assess fundamental musical achievement in tonal and rhythmic aural perception and in reading and writing tonal and rhythmic notation...ITML results serve as an objective aid to the teacher in meeting the individual musical needs of all students as a functional part of group instruction” (p. 1). The test comprises six levels of progression, concerned with tonal concepts and rhythmic concepts, respectively. For each level, there are three subtests for each category: aural perception, reading recognition, and notational understanding (Gordon, 1970; Boyle, 1973; Sidnell, 1973; Schleuter, 1974). The required responses to the tonal concept portion of the test are either aural or reading recognition by differentiating between patterns or the identification of patterns, respectively. The *notational understanding* response involves the subject completing tones within a pattern in order to correspond with the presented aural pattern (Boyle, 1973). Summarized by Schleuter (1974), subjects are required to identify the meter of an item, "to recognize correct notation for tonal or rhythmic patterns, and to take limited dictation by completing items to fit what is heard" (p. 260). For music reading, specifically, Levels 1 and 2 of the test require the subjects to determine whether a notated melody is the same or different from what the performed stimuli, and whether the notated passage has been correctly performed. At Level 6 of the test, the reading recognition portion includes chordal notation (Shuter-Dyson & Gabriel, 1981). Grutzmacher (1985) describes the reading recognition portion of the Level 2 test, which was used in her study: "the Reading Recognition subtest (ITML T2) requires the students to listen to major and minor tonal patterns while viewing tonal pattern notation and then decide (yes, no, or in doubt) whether or not what is sounded is the same as the notation" (p. 73). According to Young (1976), the best subtests of the ITML for music reading are the tonal reading recognition and rhythm reading recognition portions. Palmer (1976) also adapted the rhythm reading recognition Level 1 subtest as part of a rhythm reading achievement test. However, the reading recognition subtests have been noted as the least reliable of the measures for the ITML (Shuter-Dyson & Gabriel, 1981).

In Boyle's (1973) review of the test, he relays its purposes as follows: for the diagnosis of strengths and limitations of musical achievement, to compare one's results with the Musical Aptitude Profile (Gordon, 1965), to assess the evolution of music literacy and rhythmic and tonal aural perception for individuals, and to compare one's results with others. Young (1976) states that the ITML measures "the direct results of teaching" (p. 97). Boyle (1973) states that the benefits of administering the ITML are its separation of the key elements of music: pitch and rhythm, the increasing levels of difficulty, and the differing means of response. The ITML is found to be reliable and valid for its objectives (Young, 1976).

3.1.2. Music Reading Tests

Watkins (1942) stated: “music educators and research workers in music have long needed objective measures of instrumental achievement” (p. 3). Watkins divided achievement tests into those measuring music symbol knowledge, and those measuring performance. Music reading is commonly measured by an achievement test. In a study of sight-singing performance by achievement testing, Demorest (1998) concluded that evaluating subjects individually bettered individual performance in the group setting. He writes, “testing is not only providing information on student progress, it is helping students to be better readers, perhaps by motivating them to spend more time practicing” (p. 189). Scoring methods are central to consistent and reliable testing. Watkins (1942) states that an error must be “sufficiently noticeable” (p. 44) for all evaluators.

Historically, Mosher (1925) sought to establish “a criterion for singing achievement” (p. 15). For this portion of a larger test battery, Mosher included a series of 12 increasingly difficult sight-singing exercises. Two pilot experiments established a scoring method, which concerned tonal and rhythmic errors. Mosher and his assistants decided that they would mark a ‘-’ above or below a measure if there was a tonal or rhythmic error, respectively. If there were no errors, no mark was written on the measure. Therefore, a complete measure was the scoring unit, and Mosher found a high level of reliability among evaluators using this system. This sight singing test has not been used or adapted since; however, it is one of the early tests created for music reading achievement. An error classification method by Stelzer (1938) stated the need for a sight reading scale for the organ, identifying a "lack of reliable and valid criteria of likeability and of difficulty" (p. 35). He acknowledged errors as those that affected "notes, time, and rhythm" (p. 38). Simultaneous performance of notes, such as the playing of a chord, was scored as one unit, with the timing of the performed chord providing a second score. A true 'note method' of scoring would entail assessing each note individually. Thus, Stelzer employed a variant of the 'note method' for his unit of scoring, where there were two criterions to assess for a vertical group of notes (Watkins, 1942). Errors of fingering, pedaling, organ registration, or expression were not included. Stelzer (1938) concluded, "The number of errors made in time and in pitch is a reliable and valid criterion in grading the difficulty of reading organ music at sight. In general, if the number of errors in reading exceeds one for each measure, the composition is still too difficult for reading purposes" (p. 43).

Available music reading studies employ standard, published tests or author-composed tests. The studies present a variety of production methods, with the primary delineation being those that use sight reading or rehearsed reading tasks, or those that use sight-singing tasks. The Watkins-Farnum Performance Scale (Watkins & Farnum, 1954) is the only standardized, published performance achievement test for instrumental sight reading and rehearsed reading, while the Belwin-Mills Sight Singing Achievement Test is the only published performance achievement test for sight-singing.

3.1.2.1. Sight Reading or Rehearsed Reading

Standard Test – The Watkins-Farnum Performance Scale (WFPS)

The Watkins-Farnum Performance Scale (WFPS) began as a sight-reading test for the cornet or trumpet by Watkins (1942), and was influenced by the work of Mosher (1925) and Stelzer (1938). After a thorough examination of instrumental method resources, the scale was revised by Watkins and Farnum (1954) to include additional instruments including the baritone, clarinet, bass clarinet, saxophone, oboe, flute, french horn, trombone, bassoon, tuba, and snare drum; thus, its purpose is as a performance achievement test for band instruments (Scripp, 1995; Lehmann & Ericsson, 1993).

Watkins (1942) used a complete measure as the scoring unit, and justifies this decision:

> If the note is used as the scoring unit, certain difficulties become manifest. First, there is a tendency for any mistake so to upset the students that for the next few notes immediately following the error there is a greater chance for other errors...A whole note sustained on a high pitch is obviously very different as a response, both qualitatively and quantitatively, from a single note in a sixteenth note run and on a low pitch. Here, the unit of scoring is uneven; no two units are the same in any sense. A good measuring instrument should use equal units of measure. If the measure is used as the scoring unit at least one factor is equal from measure to measure (p. 43).

If there was an error found at any point in a measure, the entire measure was recorded as incorrect. Watkins classified the following errors, which were adopted for the final WFPS (Watkins & Farnum, 1954): errors of pitch, errors of time, change of tempo errors, expression errors, slur errors, rest errors, holds and pauses, and failure to observe repeats.

Legore (1981) describes the scale as a standardized test for “individualized administration to instrumental students and requires performance on the part of the respondent” (p. 13). The WFPS includes a series of 14 progressively difficult melodic tasks (Watkins & Farnum, 1954; Legore, 1981; Scripp, 1995). According to Miller (1988), the WFPS measures both sight-reading and rehearsed reading, with Watkins and Farnum (1954) naming the latter “practiced performance” (p. 4). There are two forms of the test: A and B, which correlate at 0.953 for sight reading and 0.947 for practiced performance. Scripp (1995) supports his use of the WFPS with the fact that it is “an accepted standard measure of sightreading for incoming instrumentalists” (p. 93), that it can be used for “test-retest measures of progress” (Watkins & Farnum, 1954, p. 9), and that its design has a foundation in earlier music reading tests.

During the WFPS, the subject hears a metronome to establish the tempo for each exercise, which stops as the subject begins to play. The subject plays each exercise as written to the best of their ability, and is required to observe note values and other markings. There is a brief pause between exercises. To score the WFPS, each measure of music is the scoring unit, for one point. The following are considered: errors of pitch, time, expression, slurs, rests, pauses, and repeats. Pitch errors comprise added or omitted tones, or an incorrect pitch played. Time errors include the misplay of note duration, as well as increasing or decreasing the tempo. McPherson (1994; 1995) discusses his reasoning for choosing the WFPS to measure sight reading in his study: “An important reason for choosing the WFPS was the scoring method, which focuses an evaluator's attention of reading ability to the exclusion of all other factors. Assessment for the WFPS is based on accuracy of performance in the categories of pitch, rhythm, slurring/articulation, tempo, expression, pause/fermata, and repeats...although in the WFPS concepts of tone quality, intonation, musical style and interpretation, phrasing, and ability to keep in time with an accompaniment are not considered in the overall scoring" (McPherson, 1994, p. 221). However, there is an alternate scoring method offered in the manual by Watkins and Farnum (1954), which records specific faults, as used with adaption by Miller (1988): "errors of pitch, rhythm, and expression were recorded independent of one another...measures without error in any one element could then be totaled. This procedure resulted in scores representing the number of measures correct in (1) pitch, (2) rhythm, and (3) expression...by counting only completely correct measures, a fourth score was derived (hereafter called the normal score)" (p. 105-106).

The WFPS is the only published test for music reading assessment and is thus common in literature, administered in, but not limited to, the studies of Boyle (1968), MacKnight (1975), Elliott (1982a; 1982b), Thompson (1985), Miller (1988), McPherson (1994; 1995), McPherson, Bailey, and Sinclair (1997), Eastlund Gromko (2004), and Hayward and Eastlund Gromko (2009). An adaptation of the Watkins Farnum Performance Scale for stimuli has been used by Scripp (1995) and an adapted scoring procedure from the original test has been used by Grutzmacher (1985; 1987), Lemay (2008), and by Pike and Carter (2010). The WFPS has been used largely with wind instruments, but has been adapted for string instruments (Farnum, 1969; Comeau, 2009), used for rhythm reading only (Boyle, 1968), and adapted for the piano (Lemay, 2008;

Pike & Carter, 2010). WFPS assessment comprising wind instruments has an average of 77.4 subjects per study, and the subjects are usually high school- or university-aged. For the adaptation of the WFPS for the piano, Lemay's 2008 study had eight young subjects, and Pike and Carter's 2010 study, where only the scoring was based on the WFPS, had 43 adult subjects.

Author-Composed Tests

Many sight reading or rehearsed reading tests are author-constructed for a specific purpose, often because the researcher feels the measure is lacking. For example, recent studies have described scoring methods that are employed for the purposes of a specific research study. These tests are rarely used in other studies, and are lacking in standardization, though there are exceptions where a researcher uses a previous test or adapts it. For example, Gilman, Underwood, and Morehen (2002) used the scoring system developed by Gilman (2000), Elliott (1982a) used the rhythm reading test constructed by Boyle (1968), and Micheletti (1980) employed the Shake Piano Sight Reading Test. The author-composed music reading tests are described in their study descriptions, which follow.

3.1.2.2. Sight-Singing

Standard Test – The Belwin-Mills Singing Achievement Test

The Belwin-Mills Singing Achievement Test was commissioned by the Belwin Publishing Company and created by Bowles in 1971 (Lemay, 2008). It is a published test, though it is no longer available commercially. The test is designed for students in elementary school through university, though the author proposes that it can also be employed for instrumental sight reading (Lemay 2008; Scripp, 1995). Subjects are to sing an exercise at a set tempo, and errors include "any judged error" (Scripp, 1995, p. 103). The scoring unit is the complete measure; therefore a final score is calculated from the measures without errors (Scripp, 1995; Henry, 2015). According to Lemay (2008), the Belwin-Mills Singing Achievement Test has been criticized due to a lack of standardization and reliability. Colwell (1991) states, "the test was not rigorously developed and no purpose is stated" (p. 254). Goolsby (1994a) used this test to measure sight reading achievement; however, the procedure for the administration and scoring of the test is not described. What is known is that the test was used to select 24 vocalists from an initial volunteer group. Two experimental groups were then created from the highest 12 and lowest 12 scores, and the selected subjects participated in the complete experiment. Goolsby's study is reviewed with further detail concerning eye-movement measurement in Chapter 4.

Author-Composed Tests

A number of sight-singing tests are also author-constructed to fulfill particular research objectives. These tests have been rarely used and lack standardization, though there are exceptions where a researcher uses a previous test or adapts it. Elliott (1982a) used the sight-singing test by Thostenson (1967), Goolsby (1994a) used the Belwin-Mills Singing Achievement Test by Bowles (1971), and Killian & Henry (2005) used the Vocal Sight Reading Inventory tool constructed by Henry (2001). The test may also be used by the same researcher in subsequent studies (Henry, 2004). The author-composed music reading tests are described in their study descriptions, which follow.

3.2. Music Reading Studies Using Tests

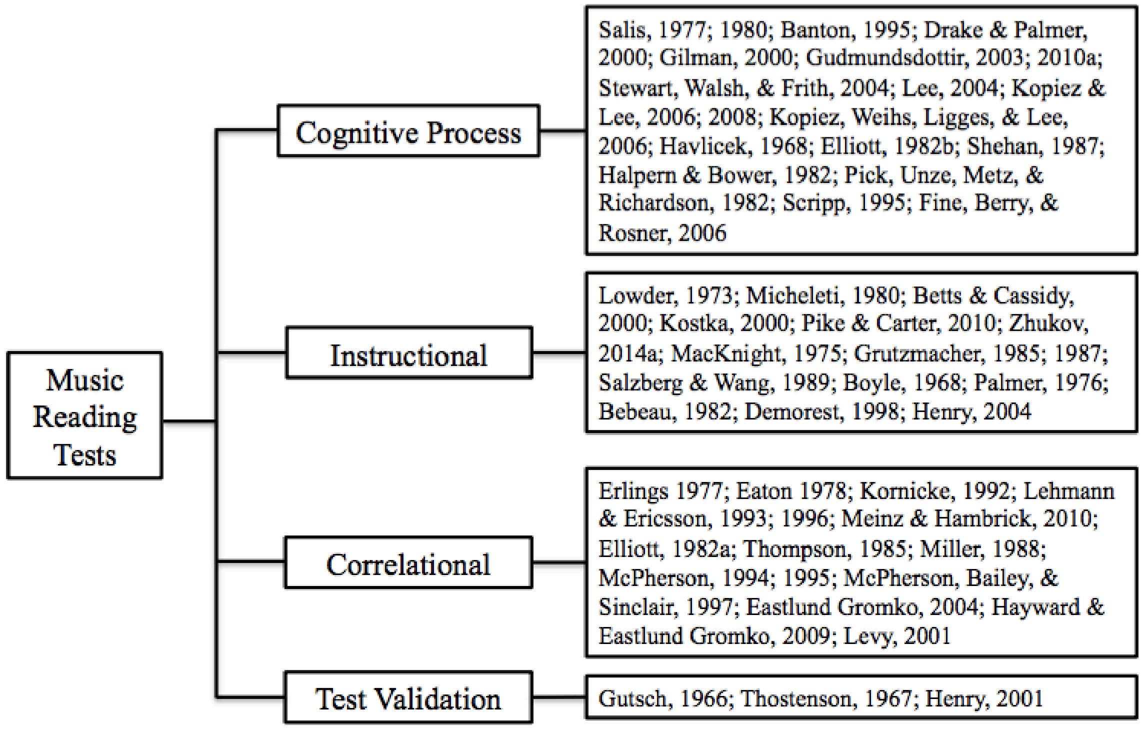


Figure 2. *Music reading tests, subcategories, and studies.*

The following studies are classified broadly as music reading tests, and fall under one of four categories: cognitive process studies, instructional studies, correlational studies, or test validation studies. Cognitive process studies comprise research concerning tasks that determine how one comprehends aspects of music reading. Instructional studies comprise research of the potential educational benefits provided by training and comparing test results between pre- and post-tests or between subject groups. Correlational studies examine components and correlates of music reading ability to determine which influence achievement. Finally, test validation studies are administered to a group or groups of subjects in order to alter the test, validate the construction, or to collect specific data.

3.2.1. Cognitive Process Studies

Piano Production

Salis, 1977; 1980	
Type	Cognitive Process
Test Construction	Author-composed
Production Method	Piano
Participants	26, university keyboard students, professional musicians, accompanists
Stimuli	45 experts of standard piano repertoire from 5 styles
Pre- and Post-Tests	No
Control Group	No
Scoring System	Expert, author-scored

Salis (1977; 1980) identified and evaluated "some cognitive aspects of the various senses used in sight reading" (p. 2), with particular focus on visual perception with advanced piano reading. To classify the subjects as either *good* or *poor* readers, the first phase of the study involved a Music Reading Test (MRT), for which Salis categorized errors. There were two large categories for error scoring: tonal and rhythmic. Tonal errors comprised omitted notes or accidentals (according to the key signature), incorrect pitches, added notes, and misread clefs. Rhythmic errors comprised incorrectly played note values, repeated notes, omitted ties, and poor synchronization of the hands. There were separate total scores for pitch and rhythm, with one point scored for each type of error (Salis, 1977; Lemay, 2008). Salis (1977) writes, "no judgment about the artistic rendition of the excerpts is included in this definition" (p. 47). The subjects were either university pianists or professional musicians, including accompanists. The stimuli for the MRT consisted of 45 brief excerpts from high-level piano repertoire. Prior to the experiment, Salis and three musicians labeled each excerpt in terms of ease of reading (*easy*, *medium*, or *difficult*). The subjects were required to perform each excerpt after 30 seconds of preview. The subjects were permitted to choose their performance tempo and were not required to play embellishments, such as trills. The performances were recorded, and Salis scored the medium and difficult excerpts according to the error classification method, with tonal and rhythmic errors. Salis also measured total time to complete the MRT, and the number of errors committed per minute. *Poor* readers, who scored below the median, committed frequent errors, both tonal and rhythmic, and performed at a slower tempo. *Good* readers were those who scored above the median. Salis did not find a significant difference between the subject groups with the measure of errors per minute. An objective of the study was to examine select subskills of music reading, including musical memory and rhythm, IQ scores of verbal and performance ability, digit span, block design and picture arrangement, the ability to read text, and mental imagery. These test scores were correlated with the Music Reading Test, and Salis found modest correlations between music reading and word reading, IQ, and musical memory.

Banton, 1995	
Type	Cognitive Process
Test Construction	Author-composed
Production Method	Piano
Participants	15, who had achieved ABRSM Grade 6 or above and regularly practicing
Stimuli	Selections from Grade 5 ABRSM repertoire; one for each condition
Pre- and Post-Tests	No
Control Group	No
Scoring System	Technology, with additional author-scoring

Banton (1995) designed a piano sight reading test to measure the function of visual and auditory feedback. The subjects sight read three excerpts under three conditions: natural sight reading; sight reading with a wooden screen that impeded the view of the keyboard and hands; and sight reading with the output speakers silenced, or a removal of auditory feedback. A different 10-12-measure stimulus was used for each condition, with all excerpts chosen from the Grade 5 repertoire for examinations (Associated Board of Music Sight-Reading Examinations, 1985). For the sight reading task, subjects viewed each piece for one minute and were encouraged to keep a consistent tempo, avoid repeating portions of the excerpt, and to continue playing. After the subjects completed the task, they rated their performances before and after hearing their recording. Subjects also reported on their musical training and involvement in a separate questionnaire, which aided in naming each subject to three specific measures: *general musical experience*, *frequency of sight reading practice*, and *sight reading ability*. The first measure was related to the subject's last achieved score on an examination. The second measure was related to "self-reported frequency of sight-reading practice" (p. 8), and the final measure was related to the subject's score for the natural sight reading condition. For the sight reading task during the experiment, there were four broad categories for error scoring: total errors, melodic errors, rhythmic errors, and specific errors. There were 10 subcategories: errors of accidentals, step errors, leap errors, omitted notes, note corrections, incorrect melodic movement based on assumption, tied notes played separately, false starts, repeats, errors of note-attack or note-release. A computer notated the recording of the performance, which was evaluated against the original notation. From this, errors were classified by the experimental condition, specific measure, and skill level. Banton found that visual feedback is necessary for accurate movements, and its absence affected all subjects. However, there was no effect when there was an absence of auditory feedback, though Banton suggests that it aids in tracking performance. The author also found that subjects with musical experience and higher reading ability influence the difference “between vital and non-vital errors” (p. 15).

Drake & Palmer, 2000	
Type	Cognitive Process
Test Construction	Author-composed
Production Method	Piano
Participants	60, 48 were children aged 7-16, and 12 were adults. All classified into group by reading level.
Stimuli	Author-composed: 3 short Western tonal pieces
Pre- and Post-Tests	No
Control Group	No
Scoring System	Technology

Drake & Palmer (2000) used a piano sight reading test to examine short- and long-term learning and their effects on the brain. For this study, there was a large group of subjects between the ages of 7 and 16, and a smaller group of adult subjects. Before the reading test, the subjects participated in an interview and performed a piece of their choice on the piano. The stimuli comprised three author-composed, Western tonal pieces. The first and second pieces were eight measures in length. The first piece had a single melodic line, while the second piece had a melodic line with two accompanying voices and a chordal accompaniment. The final piece was based on a folk tune and had melody with two accompanying voices, in 16-measure sections and in rounded binary form. The subjects began by performing the least-difficult stimulus (with the exception of the adult subjects and the advanced child subjects), with a maximum of five attempts. However, if they performed the stimulus twice without error, they continued to the next increasingly difficult stimulus. This procedure was used to determine the suitable reading level for each subject. After determining the suitable stimulus, the subject performed it five times, followed by the completion of an irrelevant task. The subjects played the piece another time, which was recorded and replayed. They were then asked to focus on one aspect of the score, though this data was not reported. With their required focus, they performed the piece five additional times. From the collected data, the subjects were assigned to one of five groups, depending on their age and sight reading ability. Drake and Palmer describe the scoring method: “Errors were identified by computer and coded according to a system similar to those used with speech errors...Five types of errors were identified in the performances: pitch errors, duration errors, pitch/duration combinations, corrections, and pauses” (p. 9). Pitch errors included errors of type, source, size, and movement. However, “multiple pitch errors whose onsets occurred within a small temporal window...were coded as chord error size; single-note errors were coded as note error size” (p. 10). Duration errors included any duration that did not agree with the score. Drake and Palmer define duration as, “the interonset interval defined by two successive events” (p. 11). Duration errors were also classified by *beat disruption.* The authors offer the example of a quarter note played as a dotted quarter note as being beat-disruptive. *Learning errors* were also coded and were described as consistent errors in the “same event location” (p. 11) during four or more of the first group of performances. “Two musician listeners” (p. 9) confirmed the computer software coding procedure by marking a transcription of each performance. There were infrequent situations where the latter coding system was used in place of the software system.

Drake and Palmer found that practice and proficiency resulted in increased and consistent tempo, and a more accurate performance. With increased skill and practice, there were more anticipatory errors; however, these subjects were also intentional with regards to planning during performance. In general, there were more combined pitch and duration errors than anticipated; more learning errors of duration than of pitch, and all subjects displayed a focus on timing over pitch accuracy.

Gilman, 2000	
Type	Cognitive Process
Test Construction	Author-composed
Production Method	Piano
Participants	24, 16 at Grade 8, 5 at Grade 6, and 3 at Grade 5
Stimuli	1-2 pieces: Grade 8 level (Guildhall School Grade Piano Sight-Reading Series)
Pre- and Post-Tests	No
Control Group	No
Scoring System	Expert, author-scored, compared with a second evaluator's ratings

Gilman (2000) measured the sight reading ability of participants in the 10 experiments that comprise her study. The subjects were required to read one selected piece, which had been chosen from "specimen sight-reading tests" (p. 43), with a metronome played throughout the performance. The scoring procedure followed the author-designed *Scoring Algorithm*. Gilman writes, "it was considered more beneficial to devise an objective scoring algorithm based on pitch and rhythm alone, that would correlate highly with the ratings of a trained expert" (p. 43). The Scoring Algorithm has three categories for error classification: pitch, rhythm, and extra notes. Pitch and rhythm errors were scored separately, with each note played receiving one point for each dimension. However, one twelfth of a point was deducted for each semitone from the original note, if played in error. Similarly, one quarter of a point was deducted for each sixteenth beat from the original beat of either note or rest, if played in error. For extra notes or rests, the durations were added and points deducted accordingly. Finally, pitch and rhythm scores are combined and divided by the highest achievable grade. Gilman offers two reasons for developing the Scoring Algorithm, or 'graded points system':

This accounts for the likelihood that in a piano exam, pianists would be awarded

> more points for performing an entire piece with inaccuracies than for only attempting part of the piece but performing it correctly. Second, the graded points system accounts for situations in which pianists read a note correctly and attempt to play the correct note but simply slip or make a slight error in motor planning. In situations like this, or in situations where the pianist simply forgets to perform a sharp or flat, then the error made will be relatively small...and will not necessarily result from poor sight-reading ability. In contrast, in situations where a note is actually read wrongly, the pianist is more likely to make a larger error, such as a major or minor third...This type of error is likely to be more indicative of sight-reading ability rather than general performance ability. Therefore, the graded scoring system is superior to an absolute scoring system as it takes into account the size of the errors made as well as the number of errors made (p. 45-46).

To assess the reliability of the error assessment method, Gilman compared the algorithm scoring procedure with expert scoring done by an "expert piano teacher...who had had several years of experience as an Associated Board examiner" (p. 48). The method was found to be highly reliable. As well as being used for Gilman's 2000 study, Lemay (2008) reports that the algorithm has been used by Gilman, Underwood, and Morehen (2002).

Prior to the main experiments, and in order to test the reliability of the Scoring Algorithm, 24 subjects participated in a sight reading task. The stimuli comprised two higher-level pieces from a sight reading method book. One was in 4/4 meter at a slow tempo, in the key of C Minor, and the second piece was in 4/4 meter at a faster tempo, in the key of G Minor. Subjects viewed each piece for two minutes, at which point they heard two measures of the set tempo with a metronome, and attempted the piece with the metronome. All performances were recorded. Gilman employed the Scoring Algorithm to assess the performances, and a second expert evaluated each performance for technical skill and expression. Gilman found reliability between the two scoring procedures, and reliability among the expert scores. She also found that expression correlated with technical skill, and remarks that this is consistent with the view that technical proficiency in sight reading allows attention to be directed to the expressive component. An interesting finding was that the data "suggests that an objective measure that can only assess technical ability would be a sufficient measure of sight-reading ability" (p. 50).

Gudmundsdottir, 2003; 2010a	
Type	Cognitive Process
Test Construction	Author-composed
Production Method	Piano
Participants	35, aged 6-13 in 2nd or 3rd year of piano study
Stimuli	Author-composed: 3 simple pieces in C Major
Pre- and Post-Tests	No
Control Group	No
Scoring System	Expert, author-scored

The objective of Gudmundsdottir's (2003; 2010a) study was to identify and categorize children's music reading errors. Stimuli consisted of three author-composed pieces with notation in both staves. The pieces were based on piano method books, and with eight measures of straightforward melody and rhythm, in the key of C Major. The first piece had alternating notation, while the latter two pieces had simultaneous notation. The 35 subjects, 6 to 13 years of age, participated in weekly piano lessons in a variety of studios, and completed the sight reading test before or after their lesson. Testing was 15 to 25 minutes in length, and the pieces were presented in a random order. The children were instructed to play each piece once, followed by a second attempt, where they were coached "to do their best to play with both hands from the beginning to the end with as few stops as possible" (p. 40). For scoring, the performances were recorded as a MIDI file and printed, with the notation *quantisized* to the nearest eighth note. Gudmundsdottir scored the printed notation, following an author-developed Error Classification procedure. According to Gudmundsdottir, "less experienced music readers are less consistent in the types and frequency of music reading errors than experienced music readers, making the analysis of their music reading performances more challenging than those of more experienced musicians" (p. 30). The classification method assessed each error by type of error, as well as its frequency and placement. There were three categories for the type of error: redundant pitches (repeated due to hesitancy), omitted pitches, and erroneous pitches. This last category of error was further divided into erroneous pitches that preserved the contour of the score, and those that did not preserve the contour. If a subject corrected the pitch after misplaying it, this was recorded as well. Three pieces were used for the stimuli and required both hands for performance, though one of the pieces alternated hands with only one hand playing at a time. The third stimulus required chordal accompaniment in the bass clef. For the scoring of this portion, Gudmundsdottir writes,

> The chord errors were initially judged according to the strictest definition possible. That is, if one pitch in the chord was incorrect the whole chord was deemed incorrect...Later a more detailed analysis was performed in order to study the nature of the chord errors made by the children in this study. In short, the first

stage of the analysis identified chords that contained errors and the second stage of the analysis identified the nature of the errors. (p. 44)

Lemay (2008) describes chord errors as categorized by redundant chords, switched chords, out-of-chord errors, within chord errors, and omitted pitch errors. Though specific rhythmic errors were not categorized, the performances were recorded, with all timing and duration data obtained from the MIDI files (Gudmundsdottir, 2003).

Gudmundsdottir found that erroneous pitches were most frequent, followed by redundant pitch errors, and the majority of errors occurred in the bass clef. The subjects would often abandon accurate rhythm or timing to perform each pitch correctly. Younger subjects committed more errors, particularly when left hand and right hand played simultaneously. Finally, Gudmundsdottir notes "the study confirmed previous findings that capacities in pitch reading skills develop gradually through childhood regardless of formal training" (2010a, p. 66).

Stewart, Walsh, & Frith, 2004	
Type	Cognitive Process
Test Construction	Author-composed
Production Method	Computer keyboard (acting as 5-key piano keyboard)
Participants	26 (exp. 1); 22 (exp. 2), all from either Royal Academy of Music or University College London
Stimuli	Each stimulus was bar of 5 notated quarter notes with numbers superimposed on note heads; In 2nd experiment, also nonmusical stimuli
Pre- and Post-Tests	No
Control Group	No
Scoring System	Technology

Stewart, Walsh, and Frith (2004) exhibited the automaticity of music notation with pianists by using two author-designed musical Stroop tasks. These were adapted from a language Stroop task, which is "one of the best known paradigms in cognitive psychology; it is also among cognitive psychology's most used contributions to clinical psychology" (MacLeod, 2005, p. 17). Stroop (1935) asked subjects to read the name of a colour presented in a different colour than the word, or the word *red* printed in the colour green. Essentially, there were two stimuli presented at the same time: a word and a colour. Stroop measured interference by timing the subjects as they read the names of colours in a black ink versus a contrasting coloured ink. In the first experiment, Stewart and colleagues (2004) presented five quarter note pitches under five conditions: (1) No notation, (2) *congruent*, where the notation was spatially correct, (3) *incongruent*, where the notation was random and spatially incorrect, (4) *incongruent*, where the notation was random but spatially correct, and (5) *catch*, where the notation was partially congruent and partially random. The 12 musician and 14 nonmusician subjects viewed each trial twice after a series of practice trials. Though the subjects did not perform on a traditional keyboard, a computer keyboard was used where the subjects' right hand fingers rested naturally, as if playing on a keyboard. The subjects viewed a fixation point, followed by the stimulus for three or four seconds (for musicians or nonmusicians, respectively), and the return of the fixation point for one second followed by the second stimulus. The subjects were asked to disregard the notation and focus on the finger numbers, reading from left to right, for performance direction. They were required to perform as rapidly as possible, with accuracy. The authors confirmed that trained pianists automatically process musical notation, and found that the notation affected musicians when asked to perform the given finger numbers, while the same stimuli did not affect nonmusicians. There was also a "relative interference effect" (p. 189) for incongruent trials.

In the second experiment the authors employed a nonmusical Stroop task. Here, 8 musicians and 14 nonmusicians participated and the stimuli included numbers presented in different locations on the screen. There were two types of presentations: a “vertical-to-horizontal stimulus-response” (p. 190) for measuring response time with spatially compatible or incompatible mappings, and a “horizontal-to-horizontal stimulus-response” (p. 190) that acted as a control. As the subjects viewed the number, they were required to press the correct finger while ignoring the spatial location of the number. The next stimulus was presented as soon as the subject responded to the previous one, but with a brief pause after five trials. Stewart and colleagues found that musicians presented a learned response of vertical-to-horizontal stimulus-responses. This is true outside of the context of a musical score. All subjects experienced interference with the horizontal-to-horizontal task.

To score the two experiments, the performances were recorded with a software program and measured for errors, response time, and cumulative analysis of each stimulus, which was presented twice. If one or more errors were committed, the trial was rejected. For response time, there were two measures: cumulative, or the total time for each sequence, and itemized, or the time for each response in a sequence.

Lee, 2004; Kopiez & Lee, 2006; 2008; Kopiez, Weihs, Ligges, & Lee, 2006	
Type	Cognitive Process
Test Construction	Author-composed, with use of questionnaire (Lehmann & Ericsson, 1996), Raven's D Matrice (Raven, 1938), and Number connection test (Oswald & Roth, 1997)
Production Method	Piano
Participants	52, piano majors or piano graduate students for Hanover University of Music and Drama
Stimuli	Selections from Four Star Sight Reading, levels 8-10 (The Royal Conservatory of Music [Toronto, Canada]) and University of South Africa sight-reading books (UNISA)
Pre- and Post-Tests	No
Control Group	No
Scoring System	Technology

Lee (2004), Kopiez and Lee (2006; 2008) and Kopiez, Weihs, Ligges, and Lee (2006) published four studies using the same sight reading test data. The test and its measurement will be described here. The 52 subjects, all proficient pianists, sight read five pieces with pacing by a metronome. The pieces had been chosen from an exam syllabus for sight reading, and "external judges" (Kopiez & Lee, 2006, p. 100), who were professional accompanists, had predetermined the level of difficulty for each piece. For the experiment, the stimuli were ordered to increase in difficulty. A composer arranged the pieces for solo violin and piano accompaniment. There were two practice pieces, followed by the five stimuli, and subjects had 60 seconds to view each piece before playing. A metronome cued the commencement of the solo part. Kopiez and Lee (2008) write, "we decided to use the piano as our test instrument due to the availability of MIDI-based methods of recording and objective procedures of performance evaluation" (p. 44). Their criteria for evaluation were for the test to be objective, able to be repeated, able to be programmed, and "independent of the raters' reliability" (p. 48). The evaluation software counted all correct pitches within a set duration.

The objectives of Lee's (2004) study were to outline and evaluate sight reading subskills with the testing of general cognitive skills, elementary cognitive skills, and practice skills. From these tests and supplementing interviews, Lee found that the "speed trill, sight reading expertise up to 10 years of age, number connection tests and inner hearing test have the highest correlation with total sight reading performance" (p. 80). The objectives of Kopiez, Weihs, Ligges, and Lee's (2006) study were to classify reading performance, and to examine whether expertise in sight reading could predict sight reading performance. They confirmed that classified sight reading performance aided in analyzing sight reading success. Expertise is a factor in sight reading, with other fixed items, such as psychomotor speed. As a continuation of Kopiez and colleagues (2006) study, Kopiez and Lee's (2006) study intended to establish a "level-specific dynamic model of sight reading" (p. 100). In addition to the sight reading task, the authors measured 23 selected skills in the areas of general cognition, elementary cognition, and expertise-related ability. They were able to confirm that sight reading achievement is a combination of skills, including information processing, psychomotor speed, and expertise in sight reading. Kopiez and Lee (2008) also wanted to establish a sight reading model, but "a general model...which can explain observed achievement variance with a limited number of predictor variables covering many levels of information processing" (p. 44). The second objective of the study was to identify the source of predictor groups. They also measured skills of cognition and expertise, and the authors found that general cognitive abilities did not influence sight reading achievement, but that select elementary cognitive abilities were predictors of achievement, and that sight reading expertise should be acquired within a crucial learning period.

Woodwind and Brass Instrument Production

Havlicek, 1968	
Type	Cognitive Process
Test Construction	Author-composed
Production Method	Piano and Instrumental
Participants	21, university music education students
Stimuli	2 melodic etudes from Twenty Studies for Flute (Gariboldi), arranged for each instrument used
Pre- and Post-Tests	No
Control Group	No
Scoring System	Expert, unknown who and how many

Havlicek (1968) used a sight reading task to study the effect of auditory feedback. The subjects were university students and were skilled in piano, brass, woodwind, or string instruments. For the experiment, four short etudes or excerpts (for piano) were chosen and arranged for each subject, according to their instrument. The performances were recorded in real-time and played through headphones. However, two of the performances had *synchronous* feedback and two had delayed feedback. With scoring, errors consisted of stopping, hesitating, incorrect rhythm or pitch, and omitting notation. Havlicek explains, “when a series of errors occurred in consecutive elements, only one instance of error was counted for the series” (p. 312). The errors were totaled for each performance. The scorer or scorers are not clearly stated for this study. It is assumed that Havlicek was a part of the scoring procedure, if not the sole judge for each performance, as was common for a study from this period of music reading research (e.g. Sloboda, 1976b). Havlicek found that there were more errors during delayed feedback, and that these effects were consistent for all performances. The responses to delayed feedback included a slower tempo, increased stopping, increased articulation errors, and added notes.

Elliott, 1982b	
Type	Cognitive Process
Test Construction	WFPS (Watkins & Farnum, 1954)
Production Method	Instrumental
Participants	30, undergraduate music theory students
Stimuli	Form A of WFPS
Pre- and Post-Tests	No
Control Group	No
Scoring System	Expert, first scored by 2 graduate students, and also by author

Elliott (1982b) administered the entire Form A of the WFPS to wind instrumentalists at the university level, in order to ascertain, classify, and compare sight reading errors of *superior* and *poor* readers. Initially, two graduate students specializing in instrumental music education scored the WFPS according to the manual, and their scores were averaged. The researcher also scored the test independently while identifying each error. Errors were identified as errors of pitch, rhythm, expression, articulation, or other errors. Pitch errors comprised key signature errors, an incorrect harmonic, a returning accidental, an enharmonic error, and other errors. Rhythm errors comprised a pause or stop, fluctuation of tempo, a meter signature error, an incorrect metronome marking, a missed rhythm, and other errors. After collecting all data, the highest 20% of subjects were labeled *superior* readers, with the lowest 20% of subjects labeled the *poor* readers. Elliott found that the majority of errors were rhythmic, and particularly for poor readers. However, superior readers made more errors of pitch. Elliott concluded, "a significant relationship likely exists between sight-reading ability and the types of errors committed by instrumental performers" (p. 38).

Rhythm Reading Production

Shehan, 1987	
Type	Cognitive Process
Test Construction	Author-composed
Production Method	Rhythm-Reading; Percussion Instruments
Participants	49, 25 in Grade 2 and 24 in Grade 6 with minimal note-reading experience
Stimuli	Author-composed: 4 8-beat rhythm patterns
Pre- and Post-Tests	No
Control Group	No
Scoring System	Expert, author-scored

Shehan (1987) measured a rhythm-reading task by recording the number of repetitions needed to achieve a perfect rendition, with a limit of 10 trials. The subjects were early- and middle-elementary students, whose "note-reading competence was assumed to be minimal" (p. 120), due to a lack of available curriculum. The author-composed stimuli consisted of four rhythms in 4/4 meter, comprising two measures of quarter and eighth notes and a quarter rest. The rhythm patterns were presented under four conditions: *audio-rhythm*, which was played on a woodblock; *audio-mnemonics*, with mnemonic syllables matched to each subrhythm; *audio-visual*, with corresponding notation on a card; and *audio-visual-mnemonic*, or a combination. During the testing procedure, the experimenter offered a brief presentation and instructions, and gave the first rhythm pattern. The subject was required to memorize and then perform the pattern on a woodblock. The experimenter scored the number of repetitions needed for a correct performance. The data suggested that a combination of visual and aural approach might be most helpful while learning and retaining rhythm, particularly for novice musicians. However, Shehan also offers that maturity may also have an important function in the speed of learning, unrelated to the method of presentation.

Production by Transcription

Halpern & Bower, 1982	
Type	Cognitive Process
Test Construction	Author-composed
Production Method	Transcription
Participants	24 (exp. 1), musicians and nonmusicians; 12 (exp. 2), 10 years experience with music reading or no experience; 14 (exp. 3), professionals or those with no music reading experience
Stimuli	"Good", "bad", and "random" melodies comprised of 10 quarter notes each
Pre- and Post-Tests	No
Control Group	No
Scoring System	Expert, 2 "musically knowledgeable" judges (exp. 2), unknown for others

Halpern & Bower's (1982) study comprised three experiments. They examined short-term memory of the melodic structure of notation, comparing musicians with nonmusicians. The pilot and first experiment employed transcription, which, for music, is an "attempt to capture in staff notation a performance" (Latham, 2002, p. 1286). The subjects, who were either proficient in music reading or had no knowledge of musical notation, briefly viewed a stimulus and attempted to transcribe the notation from memory on a given response sheet. In the pilot experiment, the stimuli consisted of 36 melodies considered *better* or *poorer*. Twenty-two musicians and nonmusicians had rated the melodies on a nine-point scale; thus, assigning the melodies to their particular category. There were also 18 author-composed melodies, considered *random*. Subjects were distracted with either a visual or auditory task for 15 seconds when the stimulus was removed, followed by their response. Halpern and Bower found that the musicians recalled more of the notation, and that both subject groups recalled the better melodies to a greater degree than with the poorer melodies.

For the second experiment, the subjects, in small groups, viewed the same melodies as in the pilot experiment in a random order. After four seconds of viewing, the subjects were to transcribe as many of the 10 notes as possible, with unlimited response time. Halpern and Bower describe the scoring procedure: "Initially, each melody was scored on an 'absolute' basis: every *n*th note of the subject's recall was compared to the *n*th stimulus note, and the response was counted correct only if these notes were identical. In the event that the subjects produced more than 10 notes, only the first 10 notes were scored. A recall of less than 10 notes was scored as the first *n* notes of the melody" (p. 37). However, the authors remark that an absolute scoring method does not allow for recall that resembles the original melody, or captures the 'musically meaningful' (p. 37) response. They offer the example of an accurate transcription of the melodic contour, but with different notes than the original. Halpern and Bower write, "the scoring method would miss the fact that a considerable amount of information had in fact been transmitted by the response" (p. 37). Therefore, judgment scoring was also applied to each response. For this scoring method, "two musically knowledgeable judges" (p. 37) gave each response a score from 1 to 10, and "judges were allowed to use any criteria of visual and/or musical similarity they felt were relevant" (p. 38), with prior direction from the authors. Inter-judge reliability was found to be r = 0.97. Halpern and Bower chose to present data from the latter scoring system. Again, the authors found that musicians could better recall the good melodies, which demonstrates that the structure of the melody is an initial part of reading process. Nonmusicians did not differentiate between the quality of the melody for recall. Random melodies were the most difficult of the melodies to recall for all subjects.

In the final experiment, the subjects viewed a sheet of stimuli and divided each excerpt into groups, or chunks, of notes. Though good and bad melodies were presented, Halpern and Bower did not find a difference in the grouping process, though they note that musicians marked fewer groups with the good melodies.

Production with Repetition

Pick, Unze, Metz, & Richardson, 1982	
Type	Cognitive Process
Test Construction	Author-composed
Production Method	Repetition
Participants	60 (exp. 1); 44 (exp. 2), all children in Grades 1 and 2
Stimuli	6 melodies of familiar or unfamiliar tunes, with nontraditional notation on a 5-line staff
Pre- and Post-Tests	No
Control Group	No
Scoring System	Expert, but assumed to be experimenter or assistant working with the children during the trial

In two experiments, Pick, Unze, Metz, and Richardson (1982) studied children's perception and understanding of notation and pitch. Using pitched bells, early-elementary children participated in three experimental conditions. The first condition presented the bells in ascending order. The second presented the bells in a random order. The third presented bells of one pitch. Each condition also involved a familiar or unfamiliar song, and author-composed notation on a staff, which did not adhere to traditional Western notation. The authors describe the experimental method:

> The child was first shown how the hammer could strike the individual bells to produce a tone. The experimenter then explained that it was possible to learn to play a song by learning to read the music on the music stand, and the experimenter played the song at the pace the child would

> later achieve...The correspondence of the notes on the staff and the bells was initially explained only in a general way...Training began with the selection by the experimenter of the first (leftmost) note on the staff and the instruction for the child to play the bell (also selected by the experimenter) that corresponded to that note. The child struck the bell, and the experimenter selected the next note to the right and showed the child which bell to play for that note. This procedure continued until the song was completed (p. 38).

Throughout the training session, the experimenter would direct the child to the notated pitch, and they would respond with the bell. The task was complete when the child could play through the training song once without error and without prompting from the experimenter. This method was repeated with a second stimulus, labeled a *transfer* song. Though the scoring procedure is not described in detail, it is evident that the experimenter present with the subject observed the performance errors. Pick and colleagues concluded that children understand notational structure, but that pitch notation and the second condition of a random order of bells required further training within the experiment. The use of a familiar song marginally aided in learning.

The objective of the second experiment was to examine children's perception of notation when encountering pitch modifications. Pick and colleagues presented the stimuli in one of two formats: individual notes, or an entire score. When the former was presented, the experimenter used a box to conceal the remaining notes. The early-elementary subjects viewed excerpts of two familiar songs in one of the formats, and responded with pitched bells, which were labeled. There was also a condition where structure and pitch notation had been altered. Subjects were required to practice the training melody until they achieved two performances without error, followed by two performances of the transfer melody. According to Pick and colleagues, "each child received a score representing the number of repetitions of the training melody required to reach the criterion, another score representing the number of repetitions of the transfer melody required to reach criterion, and also a score representing the difference between training and transfer trials" (p. 43). Pick and colleagues found that the two formats for stimuli presentation did not affect the subjects. Learning pitch and its relationship to the keyboard provides a method of understanding, and "it can be predicted that children will perceive sooner the equivalences that are systematic" (p. 44).

Sight-Singing Production

Scripp, 1995	
Type	Cognitive Process
Test Construction	Author-composed, adapted from the WFPS (Watkins & Farnum, 1954)
Production Method	Sight-singing
Participants	20, sight-singing students from New England Conservatory
Stimuli	Adaptations from the WFPS, with guidance from conservatory sight-singing books
Pre- and Post-Tests	Longitudinal (4 data collections over 2-year period)
Control Group	No
Scoring System	Expert, graduate students, teaching assistants, and alumni with experience

In a longitudinal study to propose and evaluate a cognitive-developmental model of music literacy, Scripp (1995) adapted stimuli from the Watkins Farnum Performance Scale (Watkins & Farnum, 1954) with guidance from established conservatory sight singing tests. The final four forms of the author-composed test measured both sight-singing and instrumental sight reading. Subjects were counter-balanced, with half performing vocally and half performing instrumentally. The comprehensive scoring procedure measured errors of rhythm, rhythm figures (concerning the duration between two notes), pitch, pitch intervals, additions, and omissions. Errors of intonation, expression, and rhythm entry or duration (within a half-beat of the notation) were not measured. Scripp also measured the final pitches of a melody for rhythm or pitch error, and recovery of an error for the same two components. The judges were graduate students, teaching assistants and alumni, and "all were experienced in teaching sightsinging and in administering standardized sightsinging examinations" (p. 108). They participated in "inter-rater reliability sessions" (p. 108) and any scoring modifications were carefully examined for validity measures. As well, judges were permitted to use recording devices and to consult the recordings frequently. Scripp found reliability to be high; within 96-99%.

From the collected data, Scripp found a suggestion of "an underlying understanding of rhythm...shared across sightplaying and sightsinging, while pitch representation varies considerably, depending on the mode of production" (p. 207). Scripp was particularly concerned with how internal representations of pitch and rhythm develop, and he found that sight singing ability allows a greater differentiation between rhythm and pitch. He concluded that vocal and instrumental music reading are distinct forms of music literacy, and that music reading has three skill dimensions which are independent of each other: production, perception, and reflection.

Fine, Berry, & Rosner, 2006	
Type	Cognitive Process
Test Construction	Author-composed
Production Method	Sight-singing
Participants	22, experienced choral singers
Stimuli	Pairs of notes for interval-singing, and 4 complete Bach chorales
Pre- and Post-Tests	No
Control Group	No
Scoring System	Expert, one judge scored all data with second judge scoring a sample for reliability

Fine, Berry, and Rosner (2006) designed an interval-singing test and a sight-singing test, which were given to 22 experienced choral singers. There were three objectives for their study: (a) to examine the effects of various types of interference on sight-singing performance, (b) to look at the impact of pattern recognition and prediction, and (c) to examine the associations between sight-singing ability and familiarity and experience. The interval-singing test consisted of pairs of intervals between a major third and a major seventh, both ascending and descending. The interval-singing test required the subjects to sing the second note of a given interval rapidly and precisely after hearing the first note played by a recording. The sight-singing test consisted of four Bach chorales, which were homophonic, rhythmically simple, and between 9 and 15 measures in length. The chorales were presented under three possible conditions: Harmonic modification to make the chorale discordant, melodic modification for larger intervals, or both harmonic and melodic modification, where only the rhythm of the original chorale remained. The subjects sang the chorales at a slow tempo after hearing the first pitch for an entire measure, and were permitted to attempt each chorale a second time. Precision and speed were measured for the interval-singing test, and any incorrect intervals received no score. Correct intervals received two points, though if there was a delay of more than 1.5 seconds, only one point was given, and the authors write, “a delay indicated less developed pattern recognition, albeit some melodic ability” (p. 437). The sight-singing test was measured for pitch and hesitation only. Fine and colleagues further describe the scoring method for this test, which was scored on a note-to-note basis: “one mark was given for singing any note at the correct time...two marks for singing the correct note but with hesitation or correction; and three marks for a correct performance...Hesitation was judged subjectively...Corrections occurred when a participant initially sang an incorrect note, and then very quickly noticed and changed to the correct note, often with a swoop or glide. Each semitone was treated as a category and poor intra-category tuning was therefore ignored” (p. 437). One judge scored all performances, with a second judge scoring a random sample. Inter-judge reliability was found to be r = 0.84, which the authors claim is “sufficiently reliable” (p. 438).

Fine, Berry, and Rosner found that the interval-singing results correlated with sight-singing achievement, which suggests that accomplished singers, generally with more experience, recognize patterns. For the sight-singing, modified melody and modified harmony caused disruption in performance, which was particularly true with less-skilled subjects. In this case, the authors suggest that notational audiation, or inner hearing, is important for music reading.

3.2.2. Instructional Studies

Piano Production

Lowder, 1973	
Type	Instructional
Test Construction	Author-composed
Production Method	Piano
Participants	23, university piano class students
Stimuli	Author-composed or author-selected, similar to class piano material
Pre- and Post-Tests	No
Control Group	Yes
Scoring System	Expert, 3 experienced class piano instructors
Duration of Instruction	1 semester

Lowder (1973) constructed a music reading test to measure sight reading ability, with a focus on pitch and rhythmic accuracy, tempo, and chosen accompaniment chords. The test comprised 15 tasks and the stimuli were excerpts from pieces either author-composed or author-selected. Lowder describes the excerpts as "typical of the materials presented during the subjects' first semester" (p. 70). Only pitch accuracy data was collected and analyzed for the 23 university students in a piano class, and the test was administered at the end of their first semester. The subjects were divided into control and experimental groups, with the experimental group receiving additional training in reading skills, interval recognition, and figured bass. Eight of the original 15 sight reading tasks were appropriate for the collection of pitch data. Each performance was recorded and scored by three "experienced class piano teachers" (p. 70). Lowder found that pitch errors were often in combination with rhythm errors, were often due to omitted accidentals, and were more frequent in the bass clef. Among other recommendations, the foremost proposal is that sight reading training should require rhythmic consistency, "even at the occasional expense of pitch accuracy" (p. 73).

Micheletti, 1980	
Type	Instructional
Test Construction	Author-composed, with use of the Shake Piano Sight Reading Test (Shake, 1957)
Production Method	Piano
Participants	40, 2nd year music majors from University of Miami
Stimuli	Daily sight-reading book (vertical-method condition) or wide-ranging musical styles (ensemble approach condition), unison community songs (control group)
Pre- and Post-Tests	Yes
Control Group	Yes
Scoring System	Expert, 3 graduate students
Duration of Instruction	10 weeks

Micheletti (1980) explored two instructional methods for sight reading: the vertical method and the ensemble approach. Micheletti describes the vertical method as "the ability of the student to read music vertically in two clefs, reading from the top note down" (p. 39) and it included learning bass chords by rote, self-marking of fingerings in a score, high-contrast exercises or the downbeat printed in red to "sharpen the visual reflexes" (p. 39), and an area to record learned repertoire. The stimuli used for this method was a sight reading method book with daily exercises. The ensemble approach required group practice and performance and therefore imposed tempo constraints, as well as multiple parts in which to learn. The stimuli included a wide range of musical styles. The 40 subjects were assigned to one of two experimental groups or a control group. For 10 weeks, the experimental groups received training in their respective method, while the control group followed a basic sight reading method book. All subjects participated in pre- and post-tests for sight reading, and Micheletti employed the Form A of the Shake Piano Sight Reading Test for the pretest, and Form B of the same test for the posttest.

James Shake developed the Shake Piano Sight Reading Test in 1957, after researching difficulties in both text and music reading. There are two equivalent forms of the test, and Shake includes a 'Pre-Performance Test' (Micheletti, 1980). According to Mishra (2014), the test measures the sight reading ability of university students. Micheletti (1980) describes the test stimuli: "items on the test include examples of single line melodies as well as two, three and four part writing" (p. 19). Scoring includes errors of pitch (worth 41 points for Form A; 39 points for Form B), duration (26 points, Form A; 27 points, Form B) and tempo (15 points, Form A; 16 points, Form B), and dynamics (13 points, Form A; 14 points, Form B).

The sight reading tests were evaluated by three graduate students, and inter-judge correlation was found to be between r = 0.966 and r = 0.981. Micheletti reports that the tests "were scored by hand, tabulating correct pitch, duration, dynamic and tempo markings" (p. 38). The judges marked an 'x' through each error and totaled the correct responses.

An additional objective of Micheletti's study was to also examine the relationships between music reading and factors including years of training, amount of theory training, and specifics of university study. The subjects completed a questionnaire and the tonal imagery, rhythm imagery, and musical sensitivity portions of the Musical Aptitude Profile (Gordon, 1965). For comparison, they also performed in pre- and post-sight reading tests. Micheletti found that the experimental methods did not result in significant gain in posttest scores over the control group. However, the subjects who had participated in university theory courses showed improvement in sight reading achievement, with instrumentalists scoring the highest, particularly for the 'duration' score of the MAP test.

Betts & Cassidy, 2000	
Type	Instructional
Test Construction	Author-composed
Production Method	Piano
Participants	39, non-keyboard music majors in class piano
Stimuli	Harmonization and sight reading exercises all 8 measures; 2 total for each task
Pre- and Post-Tests	Yes
Control Group	No
Scoring System	Expert, author-scored, with a sample of posttests analyzed by 2 independent reliability observers
Duration of Instruction	1 semester

The purpose of Betts and Cassidy's (2000) study was to examine the sight reading and harmonization abilities. During pre- and post-tests, the 39 nonkeyboardist musicians participated in exercises where they added appropriate harmony to a melody, as designated by Roman numerals above the score, and sight read short pieces in a second task. For the posttest, the difficulty level was increased for each task. For the sight reading portion of the study, subjects chose their own tempo and previewed the music for one minute. During the training weeks in the interim, one experimental group used software to aid in harmony practice, while the other group used software to aid in sight reading practice. This software contained accompaniment tracks, which could be controlled for key and tempo.

All test exercises was scored for pitch and rhythm errors and each task was assigned a total possible number of points. However, subjects could end with a negative score. It is assumed that the authors score the pre- and post-tests, and 33% of the pretests were also examined for reliability purposes ($r = 0.97$). Betts and Cassidy found that all subjects improved from pretest to posttest for both tasks. They note that the less difficult exercises for each task resulted in similar scores, though the tasks required different responses. This is also true for the difficult exercises of the tasks. Betts and Cassidy discuss this finding: "Sight-reading is similar to harmonization in that more proficiency with processing chordal structure should result in better sight-reading performance" (p. 159). The authors suggest that further study should examine the correlation between the two skills. Betts and Cassidy also found that accuracy was greater with subjects' right hands. They write that the collected data contributes "to the limited recent research available in class piano pedagogy" (p. 157), but that it also helps to establish teaching methods for sight reading ability, and for harmonization ability.

Kostka, 2000	
Type	Instructional
Test Construction	Author-composed
Production Method	Piano
Participants	69, university piano class students
Stimuli	Sight Reading: 2 pieces of 8 measures; more difficult piece for posttest
Pre- and Post-Tests	Yes
Control Group	Yes
Scoring System	Expert, 2 independent experienced observers
Duration of Instruction	10 weeks

Kostka (2000) examined the relationship between sight reading ability and error detection training. The subjects, 69 music university students in piano classes, were appointed to the following experimental groups: error-detection and shadowing, only shadowing, or self-guided practice. *Shadowing* is defined as "placing one's hands on the keys while doing an abstract of the music" (p. 115), which Kostka offers as an approach to practice. Before training sessions commenced, subjects participated in a sight reading pretest. The first experimental group was trained in error-detection, where subjects listened to a piece of music while marking errors on a score. The error-detection and shadowing group practiced with their hands over a keyboard without playing, and the last group practiced sight reading with any method of their choice. After the training weeks, there was a sight reading posttest. The sight reading tests consisted of eight measures of one of two possible pieces, in 4/4 time. The latter test was at a higher level of difficulty. For both tests, the subjects viewed the piece for 15 seconds before playing at a tempo chosen by the subject. Kostka describes the error scoring procedure: “two independent observers listened to the audiotapes while using an observation form containing three error categories: notes, rhythm, and hesitations. These were coded into eight discrete units corresponding to eight measures of music with each unit containing the letters N, R, and H...Thus, errors were counted according to whether they occurred in a specific measure, rather than by total frequency” (p. 118). There were three possible errors for each measure, or a potential score from 0-24. Kostka found that all experimental groups improved in sight reading ability in terms of rhythm and accuracy, though the rate of hesitation did not improve. However, the data implies that practice in error-detection may aid in sight reading performance.

Pike & Carter, 2010	
Type	Instructional
Test Construction	Adaptation of the WFPS (Watkins & Farnum, 1962) and author-composed
Production Method	Piano
Participants	43, university piano class students
Stimuli	Adaptation of the WFPS in combination with existing class piano book and author-composed stimuli
Pre- and Post-Tests	Yes
Control Group	Yes
Scoring System	Expert, authors, with additional reliability by 2 “observers”
Duration of Instruction	12 weeks

Pike and Carter (2010) did not administer the Watkins Farnum Performance Scale in its original form, but measured piano sight reading by adapting the WFPS (Watkins & Farnum, 1962) scoring procedure to an existing class piano book, with additional author-composed stimuli. The purpose of the study was to employ rhythm or pitch techniques that aided in viewing patterns, or *chunking* of notes on a score. After 12 weeks of instruction to piano classes in a university, with the experimental groups receiving the supplemented training, the authors compared experimental and control groups. The modified scoring procedure included the assessment of right hand and left hand accuracy for rhythm and pitch, and continuity of performance; a total of five possible points for every beat of music. All pre- and post-tests were recorded by MIDI and scored by the authors, with 20% of the tests further analyzed by "two independent reliability observers" (p. 238). Reliability for pretests was found to be r = 0.98, and for posttests, r = 0.99. Pike and Carter found that though there were not significant differences between the groups, all subjects improved in sight reading ability over the term. As the control group did not have additional training with rhythm chunking techniques, the results implied that the control group focused on pitch above rhythm. The experimental group that focused on rhythm patterns improved in accuracy of rhythm and in continuity, however, pitch processing did improve from this training. The experimental group that focused on isolated pitch patterns showed improvements in rhythm, pitch, and continuity.

Zhukov, 2014a	
Type	Instructional
Test Construction	Author-composed
Production Method	Piano
Participants	100, young adult to adult, all proficient pianists
Stimuli	Rhymically-complex (rhythm program); Baroque, Classical, 20th-Century (Accompanying and Control programs)
Pre- and Post-Tests	Yes
Control Group	Yes
Scoring System	Technology
Duration of Instruction	10 weeks

Zhukov (2014a) studied the effects of three instructional methods in sight reading. The 100 subjects, who were proficient pianists ranging from young adult to adult students, were assigned to one of three experimental groups or a control group. The objectives of the *Accompanying Training Programme* were to develop movement in performance through improved eye motion, consistent tempo, and counting. This method used Baroque, Classical, and Twentieth-Century repertoire, and the subjects learned an accompaniment part, which they rehearsed weekly with a soloist partner. The objectives of the *Rhythm Training Programme*, which employed repertoire with complex rhythms, were to better comprehend and perform rhythms. For this experimental group, the subjects analyzed the rhythm of the score, followed by a performance of the score. The *Style Training Programme* used shorter pieces from Baroque and Classical repertoire and analyzed the harmony and structure of a piece before performing it. The objective of this method was to identify distinctions between the repertoire in order to locate patterns and improve predictions for each style of music. In addition to the three modes of instruction, the subjects regularly wrote in a journal about their practice. The control group followed regular instruction and participated in musical extra-curricular events as usual.

The subjects were trained in their respective methods for 10 weeks. During this time, Zhukov met with each subject three times, where the subjects performed "the materials from the preceding weeks and answered questions with regard to rhythmic and structural analyses, and approaches to practice" (p. 74). A pretest and posttest for all subjects required them to perform three excerpts in the same musical style as their respective instructional method. They were permitted to view the stimulus for one minute, and played with a metronome. There was also an exit interview, where the subjects reported any additional sight reading completed during the training period.

For scoring, Zhukov writes, "due to a lack of any other existing programme, custom-made software was developed for this study by a small team of consultants with expertise in data analysis and musicianship" (p. 75-76). This software collected recorded MIDI data, and was designed to measure all notation, regardless of whether or not the subject paused during performance. It provided numerical data on accuracy of pitch and rhythm, with a total of four scores, and also measured beat adjustment, extra notes, omitted notes, and *root means squared* (RMS) accuracy, which is defined as the "average number of timing errors per correct note played" (p. 76). Reliability and validity of the software program was confirmed by "a series of manual checks" (p. 76). Zhukov found that the accompanying group showed consistent improvement across all measures, but that the group improved the least when considering all subject groups. The other three groups improved in all measures, with some exceptions depending on the test piece, and that the control group participated in the most additional sight reading. In particular, Zhukov found that specific instruction aids in sight reading achievement, rhythm instruction aids in continuity, analysis improves pattern detection and prediction, and additional sight reading improves "post-test performance" (p. 82).

Woodwind and Brass Instrument Production

MacKnight, 1975	
Type	Instructional
Test Construction	WFPS (Watkins & Farnum, 1954), MAP (Gordon, 1965), MAT (Colwell, 1969), author-composed
Production Method	Instrumental
Participants	Approximately 90 elementary students
Stimuli	Author-composed (experimental group); *Breeze Easy* method book (control group)
Pre- and Post-Tests	Yes
Control Group	Yes
Scoring System	Expert, unknown, but assumed to be author
Duration of Instruction	32 weeks

MacKnight (1975) studied elementary-school students studying brass or woodwind instruments, though during training the students also responded vocally. The purpose of the study was to construct an instructional technique to develop the process of music reading, with a particular focus on melody. In addition to the WFPS, MacKnight administered Gordon's MAP (1965), the Colwell MAT (1969) and an author-composed Student Attitude Questionnaire. All students, divided into experimental and control groups, were instructed in the same manner, but with varying methods of introducing pitches. The experimental group learned pitch with three means of tonal pattern instruction: aural presentation, auditory-visual presentation, and auditory-visual presentation in the context of a larger phrase. This group also learned rhythm through aural or visual phrases. In contrast, the control group used a standard book method to dictate its pitch and rhythm introductions. For the WFPS portion of the study, the test was administered after 32 weeks of training. It is assumed that the author completed scoring for the WFPS, as it is stated that the author administered the test. The test was given and measured according to the manual, and reliability was determined to be r = 0.93 for select subject groups. MacKnight explains that the findings from this study "suggest that melodic training based on the auditory-visual recognition of a series of tonal patterns is an effective strategy for teaching the meaning of notation to beginning music readers" (p. 34). MacKnight concludes that pattern training is better for both music reading and auditory-visual discrimination ability than having students identify individual notes.

Grutzmacher, 1985; 1987	
Type	Instructional
Test Construction	Author-composed, with adaptation of WFPS (Watkins & Farnum, 1954), MAP (Gordon, 1965), ITML (Gordon, 1970)
Production Method	Instrumental
Participants	48, middle-elementary students
Stimuli	20 tonal patterns and *Alfred's Basic Band Method* (1977) (experimental group); Text without tonal patterns (control group)
Pre- and Post-Tests	Yes
Control Group	Yes
Scoring System	Expert, 3 experienced judges
Duration of Instruction	14 weeks

Grutzmacher (1985; 1987) designed the Melodic Sight-Reading Achievement Test (MSRAT) as part of the examination and comparison of two approaches for the sight reading development of wind instrumentalists. The test is based on the scoring procedure from the WFPS (Watkins & Farnum, 1954), though Grutzmacher did not include rhythmic data in the analysis. One approach for sight reading training was to identify notes individually, and the other was to learn tonal patterns with harmonization and vocalization. Grutzmacher asserts, "tonal pattern perception is more efficient for music reading comprehension than is the perception of individual pitches" (1985, p. 11), supporting this claim with the commonly held views that musicians perceive pitch patterns, and associate tones by their shared tonal centre.

The MSRAT was administered as a posttest after 14 weeks of training in either tonal patterns or note-identification. It contained six short author-constructed melodies with tonal and rhythm patterns from the training curriculum, and the level of difficulty was varied. The performances were recorded and assessed "by three judges not including the researcher. All three judges were Ohio state certified music teachers with a minimum of two years of public school teaching experience in instrumental music" (p. 74). Only pitches within the tonal patterns were scored with two points awarded for each correct pattern. "While each measure of notation is the unit of scoring in the WFPS, each tonal pattern is the unit of scoring in the MSRAT" (p. 75). A first error in a particular pattern resulted in a one-point deduction, and a second deduction for subsequent errors. Grutzmacher found that the experimental groups showed higher scores on the MSRAT than the control groups. Sight reading ability was improved by vocalizing with tonal patterns, and Grutzmacher suggests that sight reading instruction include curriculum to develop both aural and visual perception proficiencies.

Salzberg & Wang, 1989	
Type	Instructional
Test Construction	Author-composed
Production Method	Instrumental rhythm-reading
Participants	46, ages 8-16 with 1-7 years' experience
Stimuli	Rhythmic exercises at 10 levels of difficulty
Pre- and Post-Tests	Yes
Control Group	No
Scoring System	Expert, one of the authors and a second judge
Duration of Instruction	Unknown

Salzberg & Wang (1989) studied rhythm reading as influenced by the use of three prompts: foot tapping, counting aloud, and a combination of tapping and counting aloud. The subjects comprised 46 string students between 8 and 16 years of age. The experiment began with a pretest of rhythm sight reading, in order to place each subject at the appropriate level of difficulty. For the pretest and posttest, Salzberg and Wang constructed 10 levels of sight reading exercises, ranging from Level 1, which was notated with quarter, half, and whole notes and quarter rests, to Level 10, which was notated with the aforementioned notes and rests, as well as dotted half, eighth, dotted eighth, sixteenth notes, and half and whole rests. Each level had four rhythms of eight measures each to be played on an open string. From the pretest data, subjects were assigned to a rhythm reading level (from one to five) and received instruction and practice in each of the prompt conditions, accordingly. The authors administered the pretest for two reasons, as explained by Salzberg and Wang: "to ensure that each student would be reading...at approximately the middle of his difficulty range, and thus ensure that he would not 'peak out' and score very high...nor be reading music so difficult that no treatment would be effective; and to ensure that each of the four exercises within each level were of comparable difficulty and that the levels...were an accurate progression of difficulty" (p. 124). For the posttest, the subjects played the four exercises of their current level, in a random order, and each with a different prompt. The performances were recorded and scored note-by-note, with each quarter note beat equal to one point. All performances were evaluated at two instances by one of the researchers. A second evaluator independently scored a sample of the performances for reliability purposes, and reliability was found to have a mean of $r = 0.885$. Salzberg and Wang found that the subjects in the three lower reading levels improved in rhythm reading with the counting prompt; thus, it was effective for less-skilled musicians. The data for subjects in the two higher reading levels did not show any difference in rhythm reading between prompts.

Rhythm Reading Production

Boyle, 1968	
Type	Instructional
Test Construction	Author-composed, adapted from the WFPS (Watkins & Farnum, 1954)
Production Method	Instrumental rhythm-reading
Participants	201 (pretest); 191 (posttest), all secondary students
Stimuli	Adapted from the WFPS
Pre- and Post-Tests	Yes
Control Group	Yes
Scoring System	Expert
Duration of Instruction	14 weeks

Boyle's (1968) examination of the effect of rhythm training on rhythm sight reading, or rhythm reading, is one of the first studies to focus on this component of music reading (Miller 1988; Kornicke, 1992). He sought to learn whether rhythm training would improve music reading, with both rhythm reading and with performed music reading. Boyle writes, "many music educators believe that music reading deficiencies are in large part a result of students' inability to read and perform rhythm patterns correctly" (p. 2). Boyle used Forms A and B of the WFPS (Watkins & Farnum, 1954) as the stimuli for a rhythm reading test, by taking the exercises from the WFPS and notating them on a single pitch in the treble or bass clefs, depending on the instrument to perform the exercises, and without expression or dynamic indications. There was also a form written in percussion notation for snare drum. The subjects in Boyle's study, high school instrumentalists, completed a pretest of the sight reading measures, followed by 14 weeks of training. The control and experimental groups were instructed with the same curriculum, while the experimental group had supplemental training consisting of listening to recordings, marking a beat on a score, tapping to a beat, clapping and tapping to a beat, and playing rhythm patterns on a single pitch while tapping. After the training weeks concluded, the posttest consisted of one form of the rhythm reading test and the opposite form of the original instrumental WFPS. Three judges evaluated the subjects, and tests were scored according to the WFPS manual. Boyle describes the scoring procedure: “the rhythm sight reading test...required value judgments by the scorer. Therefore, each subject’s performances on these measures were tape recorded, randomized and then scored independently by the jurors. The mean of the three jurors’ scores for each subject constituted that subject’s score” (p. 76-77).

Boyle found that both groups improved on the rhythm reading test and the WFPS, but that the experimental group showed significant gains. The rhythm training aided in improving instrumental sight reading ability for the experimental group. Boyle concludes, “because the reading of rhythms is a necessary part of music sight reading, the correlation coefficient of .81 between students’ scores on measures of the two abilities is interpreted as supporting the view that the ability to sight read music is dependent in large part upon the ability to sight read rhythms” (p. 91). Elliott (1982a) used Boyle’s test with a slight adaptation, and Elliott clarifies that scoring errors followed the instructions for the rhythm portion of the original WFPS, with a final tally of correct measures. Boyle (1968) and Elliott (1982a) report a high reliability for this rhythm reading test.

Palmer, 1976	
Type	Instructional
Test Construction	Pretest: Rhythm reading achievement methods of Richards *Threshold to Music* and E. E. Gordon method, MAP (Gordon), MAT I & II (Colwell), ITML (Gordon) Posttest: Richards and Gordon rhythm reading achievement methods
Production Method	Rhythm-Reading (Unknown production)
Participants	136, children in Grade 4
Stimuli	*Threshold to Music* (Richards); E. E. Gordon method
Pre- and Post-Tests	Yes
Control Group	Yes
Scoring System	Expert, 3 elementary music specialists
Duration of Instruction	5 months

Palmer (1976) examined the efficacy of two rhythm reading instructional methods. The first was based on the Kodaly method (Richards, 1971), and the second was the Gordon reading method (Gordon, 1971). Six classes of early-elementary subjects were divided into control and experimental groups. The control class did not have a specific rhythm-reading program, while the experimental classes participated in the specific rhythm reading training in three classes each week for five months. The pretest included portions of the MAP (Gordon, 1965), the Colwell MAT (Colwell, 1969; 1970), and the ITML (Gordon, 1970). The posttest was an author-composed, prerecorded "performance achievement measure" (p. 111) in three parts. The *Response to Meter* portion measured synchrony to duple and triple pulses while listening to music. The *Imitation of Rhythmic Patterns* portion measured the response to an aural rhythmic pattern. The *Rhythmic Notation* portion measured the response to rhythm in a score. It is not stated as to how the subjects responded to the subtests. A five-point scale was used for scoring, and the evaluators, "three elementary music specialists" (p. 111), independently scored the performances, which were averaged for a final score. Palmer found that "a systematic approach to the development of skill in rhythm reading resulted in achievement gains" (p. 112); thus, the experimental groups significantly improved in rhythm reading ability in contrast to the control groups. Palmer also noted that written rhythm and performed rhythm are separate components of rhythm reading ability.

Bebeau, 1982	
Type	Instructional
Test Construction	Author-composed
Production Method	Rhythm-Reading with Verbal and Clapping Response
Participants	27 (exp. 1); 80 (exp. 2), all children in Grade 3
Stimuli	Author-composed rhythm patterns
Pre- and Post-Tests	Yes
Control Group	No
Scoring System	Expert, 3 experienced music teachers
Duration of Instruction	18 days

Bebeau (1982) also studied two rhythm reading instructional methods. The subjects of the first experiment were early-elementary students, divided into two experimental groups of 14 and 13 subjects, respectively. The first instructional method used associated speech and action cues to present rhythm reading. The second method was considered a traditional method, with the subjects counting pulses while clapping or playing the rhythms. Both methods introduced 12 rhythm concepts. An author-composed criterion test was used for both the pretest and posttest, and students were tested individually, responding with clapping, or clapping and speech. The test consisted of 23 author-composed rhythm patterns in 4/4 meter, spanning one to three measures in length, and excluding syncopation or tied notes. They were presented in order of increasing difficulty. In addition, there were two items of 8 and 10 measures of simple rhythm, respectively, to measure the subject's ability to maintain a consistent tempo. Bebeau writes, "three experienced music teachers participated in the development and field testing of the test and scoring procedures" (p. 112). Scoring was based on a five-point scale, evaluating accuracy of rhythm and tempo. However, the marks were weighted towards accuracy of rhythm over accuracy of tempo, "since the performer's inability to read the rhythmic pattern is most often responsible for variations in tempo" (p. 114). The judges scored the performances independently and at different times because it was obvious as to which experimental group the subject belonged. Scores were compared and reliability was found to be between r = 0.80 and r = 0.96. The scores were averaged for a final score. Bebeau found that the speech-cue group had significantly higher scores.

In the second experiment, 80 subjects participated in an identical method to that of experiment 1. Here, the scoring procedure is not described, though it is assumed that it also followed the first experiment. Bebeau additionally analyzed 30 subject's results to compare gain scores. Though the posttest comparison did not show significance, the results of the gain scores again favoured the speech-cue participants.

Sight-Singing Production

Demorest, 1998	
Type	Instructional
Test Construction	Author-composed
Production Method	Sight-singing
Participants	306, secondary choral students
Stimuli	Adapted melodies from *Music for Singing* (Ottman, 1967), *Oxford Folk Song Sight-Singing Series* (Crowe, Lawton, & Whittaker, 1961)
Pre- and Post-Tests	Yes
Control Group	Yes
Scoring System	Expert, 2 trained evaluators
Duration of Instruction	1 semester (less a month)

Demorest (1998) studied group training and individual assessment for sight singing achievement. The 306 subjects from beginning and advanced high school choirs completed a musical background questionnaire and a sight-singing task as a pretest, and continued receiving regular classroom instruction for a semester. In addition, the experimental group participated in individual testing sessions where they were to record a portion of a melody from the practice material, and received feedback on both strengths and weaknesses from their performance. This occurred periodically throughout the semester. At the posttest, all subjects completed another sight-singing task with a major melody and a minor melody, adapted from a sight-singing method book. For the sight-reading tests, Demorest scored errors of pitch, including repeated notes, and restarting a melody, and rhythm, including changes of tempo. The scoring process is further described: "All scoring was done from the tape-recorded performances using the following system. One-half point was deducted for each rhythm error, and one-half point was deducted for each pitch error. One full point was deducted for a tempo change or for a repeated note. Subjects were penalized 2 points for starting the exercise over. Each example consisted of 15 notes, so a score of 15 constituted a perfect score" (p. 184-185). Each performance was assessed by one of two trained evaluators, who were blind to the nature of the experiment. A sample of scores from each evaluator showed reliability at r = 0.97. Demorest found that individual testing might improve sight-singing achievement, as the experimental group showed significant progression from the pretest to the posttest with regard to the major melody. There was not the same effect for minor melodies.

Henry, 2004	
Type	Instructional
Test Construction	Henry Vocal Sight Reading Inventory (Henry, 2001)
Production Method	Sight-singing
Participants	67, secondary students
Stimuli	Author-composed exercises with the targeted pitch skills, with solfege syllables placed vertically according to the relative 'height' of the pitch, and 12 familiar melodies in standard notation with solfege syllables next to each note head
Pre- and Post-Tests	Yes
Control Group	No
Scoring System	Expert, author and a graduate assistant
Duration of Instruction	12 weeks

The Vocal Sight Reading Inventory (Henry, 2001) was employed in Henry's (2004) study. The scale was developed in order to assess vocal sight-singing ability, with additional objectives of allowing "formative and summative evaluation" (Henry, 2001, p. 21) and establishing "validity, reliability, and normative data" (p. 21). The VSRI is designed for individual administration for high school students; as such, another objective was that the duration of the test be compatible in the choral classroom setting. There are two forms of the test and each consists of eight major key melodies notated in the bass and treble clefs and without complex rhythms. The melodies contain embedded pitch skills, which Henry obtained from tonal music. She named seven categories in which to classify the pitch skills, or functional melodic patterns: conjunct, which are simple ascending or descending patterns; tonic; dominant; subdominant; cadential, with four authentic cadences; modulatory; and chromatic. The test itself requires subjects to listen to a chord progression of the key and the first pitch of the melody. They view and, if desired, practice the melody for 30 seconds before the chord progression and first pitch are played again, and then sing the melody. The scoring procedure applies to the pitch patterns within the melody, by codes associated with each pitch. Incorrect pitches are scored, though the final score given to the subject represents the number of correct component skills executed. For the 2001 study, Henry and a second scorer evaluated recordings of the pilot performances with the scoring method proposed by the VSRI. For the main study, Henry scored the VSRI with the proposed method and a second scorer independently scored using a note-by-note method, for inter-scorer reliability. This was found to be r = 0.97.

The objective of Henry's (2004) study was to assess the result of emphasizing particular pitch skills during choral instruction, and a portion of the VSRI was given as a pretest. The 67 high school subjects were divided into a *solfege drill* group and a *familiar melodies* group. For training, the first experimental group practiced pitch patterns using solfege notation, where the syllables were written in a pitch hierarchy. The second experimental group practiced pitch patterns within familiar melodies, which were presented with standard notation and solfege syllables. After 12 weeks of training, the complete VRSI was administered. Henry found that all subjects improved in sight-singing by practicing pitch skill patterns, regardless of their experimental group.

3.2.3. Correlational Studies

Piano Production

Erlings, 1977	
Type	Correlational
Test Construction	Author-composed, MAP (Gordon, 1965), MAT (Aliferis, 1954)
Production Method	Piano
Participants	33, university class piano students
Stimuli	Materials according to specific tests, and Author-composed with arranged German dances by Haydn
Pre- and Post-Tests	Yes
Control Group	No
Scoring System	Expert, author-scored, with sample scored by second judge

Erlings (1977) was concerned with predictive factors associated with rudimentary music reading levels and two aptitude tests: the Musical Aptitude Profile (Gordon, 1965) and the Music Achievement Test (Aliferis, 1954). These tests were administered in combination with a motor control test, a sight reading test, and information regarding the subjects' grade point average (GPA) and American college test (ACT) scores. For the piano sight reading test, which was designed for university students, there were two arranged sight reading pieces at the required level of proficiency. A pretest allowed for materials, test administration, and scoring procedures to be established. Though the scoring procedure is not explained in this publication, errors were counted, and ranged from 0 – 63.5, with a passing score between 0 and 34. As well, Erlings states that "scoring standards and procedures...anticipated many possible variables and were outlined in detail" (p. 15). Erlings scored the two subtests of sight reading during the main experiment, evaluating pitch, rhythm, tempo, and continuity. These two scores were averaged. A "qualified piano teacher" independently evaluated a sample of the tests, with an inter-judge correlation of r = 0.995. The author concluded that the sight reading test was a successful quantitative measurement tool, and that the scoring and administrative procedures could be repeated to further “objective evaluation of psychomotor keyboard skill achievement” (p. 17).

Erlings found correlations between sight reading and tonal and rhythm imagery scores of the MAP, and the melodic and rhythmic portions of the MAT, though only the melodic portion of the MAT was statistically significant. GPA scores were also correlated, though motor control and ACT scores were not. This study serves as an initial examination of keyboard reading and its subskills, and Erlings concludes, “it could be beneficial for teaching and advising purposes to investigate relationships between kinds of error made in keyboard sightreading tests and musical aptitude and achievement measurements” (p. 17).

Eaton, 1978	
Type	Correlational
Test Construction	Author-composed
Production Method	Piano
Participants	73, university-level pianists and organists, community keyboardists
Stimuli	Sight Reading: Selections from Daquin, Clementi, Dvorak, Jarnach, Memorization Skill: Hofmann, Schytte
Pre- and Post-Tests	No
Control Group	No
Scoring System	Expert, 3 judges

Eaton (1978) studied 73 pianists who were proficient in music reading. The subjects completed a reading performance task, followed by a note-naming task, which included both verbal and played responses. For the note-naming task, subjects named single notes, intervals, and triads as rapidly as possible. For the performance task, four published piano pieces were used as the stimuli. Scoring of the Sight-Reading Facility portion of the study followed the author-composed *Grading Instructions*. To record errors for the music reading portion of the test, Eaton developed a Subject Score Sheet. Lemay (2008) describes the instructions:

> Errors were divided into two categories: note errors, and rhythm and tempo errors. In the note error category, judges were instructed to circle wrong notes, notes within chords which were incorrect, missing notes, as well as the re-striking of notes or chords. All wrong notes were given one mark. In the rhythm and tempo error category, judges were instructed to give one mark for short hesitations, two marks for long hesitations, one mark for change of pulse, one mark for incorrect rhythm, one mark for failing to observe a fermata, and three marks for failing to observe the indicated tempo. (p. 21)

According to Eaton (1978), there were also spaces to record the time taken to complete the subtest, and the time-penalty for errors, for which he used the following formula (p. 68):

$$\frac{\text{Total Errors}}{\text{Total Possible Answers}} \times \frac{\text{Elapsed Time}}{1} = \text{Penalty}$$

Three judges scored the tasks. In addition to being faculty members of a university, Eaton describes their qualifications: “two of the judges...hold master’s degrees in piano. The third judge...is a performing pianist as well as an instructor in keyboard harmony" (p. 70). Over three sessions, the judges were trained in the Grading Instructions by hearing good, average, and poor samples of performances, and completed scoring. Inter-judge reliability was found to be between r = 0.989 and r = 0.994. Using a score sheet, Eaton recorded the judges' error scores, calculated an average, and analyzed the total scores for note errors, rhythm and tempo errors, and sight-reading errors.

Eaton also studied the relationships between keyboard training and experience with select abilities for proficient musicians. During the individual tests, subjects offered background information and then completed tests of sight reading, note-reading with verbal and performed responses, psychomotor skill with visual and oral commands, and memorization skill with practice and visual analysis. Eaton found that the correlation of significance was between sight reading and psychomotor ability. Lesser correlates were note-reading and performance experience.

Kornicke, 1992	
Type	Correlational
Test Construction	Myers-Briggs Type Indicator (MBTI) (Myers & Briggs, 1976), Author-composed sight-reading test (SRAT), Hidden Figures Test (HFT) (French, Eckstrom, & Price, 1962)
Production Method	Piano
Participants	73, university students and professors
Stimuli	Sight Reading: Author-selected from available concert repertoire, Aural Imagery Test: Recorded chords on acoustic piano.
Pre- and Post-Tests	No
Control Group	No
Scoring System	Expert, doctoral-level students in piano performance

Kornicke (1992) assessed piano sight reading by looking at a variety of predictors for achievement, to determine which might be the best predictor of success in reading. To measure the sight reading ability of the subjects, who were undergraduate, graduate, or professional pianists, Kornicke developed the Sight-Reading Performance Scale (SRPS). The scale comprised five musical excerpts, selected from concert repertoire. The scale was part of a larger test, called the Sight-Reading Achievement Test. This involved an interview concerning sight reading experience, followed by the SRPS. Subjects played all five excerpts at their own tempo and with the goal of performing the excerpt. Each excerpt was recorded and then scored by 32 items on a five-point scale (*seldom* to *usually*). The majority of the items were consistent for each musical excerpt, such as "appropriate shaping of melodic line...appropriate dynamic nuance in phrases" (p. 252). There were also eight items specific to a particular excerpt, such as "staccato notes played evenly " (p. 252) or "tempo irregularities occurred in short sections of piece" (p. 255). To support this itemized scoring system, Kornicke writes, "by scoring the sight-reading achievement test for interpretative aspects, a dimension was utilized for examining sight-reading that more accurately fits the definition of sight- reading as the performance of a piece of music without prior practice" (p. 207). The SRAT was assessed by three "doctoral level students who had a completed Masters degree in piano performance" (p. 109). The judges had previously studied the items on the score sheet and participated in a training session with recordings from the pilot study, and inter-judge reliability was found to be r = 0.99.

As well as the sight reading task, subjects completed an experience questionnaire, and tests of aural imagery, hidden figures (for field dependence and independence), locus of control, and the Myers-Briggs Type Indicator (MBTI) (Myers & Briggs, 1976). Among the many findings, the key conclusions from the study were that sight reading ability correlated with aural imagery, sight reading experience, and field independence, as the “best linear combination of predictor variables” (p. 183), and that the predictor factors which differentiated the successful and less-successful subjects were sight reading experience, field independence, aural imagery, and the *thinking* score from the MBTI. Kornicke notes that sight reading experience is a “complex variable” (p. 154) and includes amount, frequency, and array of experiences. The author also found a differentiation between sight reading and performance, and states that this reveals that the ability “involves skills that are not likely transferred from the sole act of practicing and/or performing literature” (p. 155).

Lehmann & Ericsson, 1993	
Type	Correlational
Test Construction	Author-composed
Production Method	Piano
Participants	16, performance and accompanying students from Florida State University School of Music
Stimuli	Moderately difficult pieces for solo instrument and accompaniment (Schubert, Blavet, Molique). Each excerpt was two pages in length.
Pre- and Post-Tests	No
Control Group	No
Scoring System	Technology

Lehmann & Ericsson (1993) used excerpts from standard repertoire to measure sight reading achievement; to better understand differences of ability. All subjects were undergraduate pianists, with eight focused on accompanying and eight focused on performance. The stimuli were arranged for a solo instrument with piano accompaniment, thus the subjects played two 2-page excerpts with a prerecorded track. This task paradigm imposes a consistent tempo (Kopiez et al., 2006). The subjects were required to play each excerpt four times: two times with a metronome and two times without the metronome or the solo track. Lehmann and Ericsson write, “the most obvious problem in measuring instrumental skills is that an individual's rehearsed performance of specific piece will vary greatly as function of how long the piece has been practiced and on how complex its musical structure is" (p. 183). This method resulted in a measure of both sight reading performance and of rehearsed performance for each excerpt. Music editing software allowed the collected data to be recorded and transcribed as notation to the nearest sixteenth note, which was scored for pitch and timing. Only correct items were scored, and scoring details are as follows: if "the note had the required pitch, the actual onset coincided with the prescribed onset, or the note was notated as an associated grace note" (p. 189). The first and last lines of the excerpt were not included in scoring. As well, "two experienced jurors" (p. 189) evaluated the performances using a six-point scale for their general sense of the performance, the technical component, and expression. Due to the subjective nature of this evaluation, Lehmann and Ericsson state, "the raters were presented with an example for very good and a very poor performance at the beginning of each session" (p. 189). Scoring of the software transcription was compared with the jurors' scores, and the two were found to be reliable (r = 0.88), thus validating the sight reading algorithm used with the software program. The authors found that experience in accompanying positively influenced sight reading achievement, and that at a high level of proficiency, pianists are more likely to vary due to specialized skills.

Lehmann & Ericsson, 1996	
Type	Correlational
Test Construction	Author-composed
Production Method	Piano
Participants	16, performance and accompanying students from Florida State University School of Music
Stimuli	Author-selected from Lehmann & Ericsson (1993)
Pre- and Post-Tests	No
Control Group	No
Scoring System	Technology

Lehmann and Ericsson's (1996) study was a continuation of Lehmann and Ericsson (1993). The objective of this study was to determine the cause for differences in sight reading skill. The same 16 subjects participated in both studies, and completed three tasks with three separate stimuli. The first task was a recall exercise, where a portion of a previously seen excerpt (from Lehmann and Ericsson, 1993) was presented, but with select notation omitted. The subjects were asked to play the excerpt with a metronome and perform as many of the missing notes as possible. The second task involved a new excerpt, and the subjects were required to improvise at any of the omitted portions. The final task was considered a *leap* task, and the stimulus comprised large leaps for both left and right hands. For this exercise, the subjects learned the chord structure, played the excerpt once with the metronome, and again with the metronome and goggles to limit peripheral sight. All performances were recorded and a software program notated each. For the first tasks, the excerpts were compared to the same sight read excerpts from the previous study. Only the matching notes between the two performances were scored, and only for selected portions. Correct notes were scored for the final task, with select notes from each attempt totaled. Lehmann and Ericsson (1996) refer to the previous study for the scoring procedure, thus, it is assumed that the scoring methods were similar.

The authors found that predictors of sight reading ability included training in a specific area, such as accompanying, and knowledge of that area, such as common accompaniment repertoire. The authors conclude, "individual differences in sight-reading ability...do not seem to reflect innate music talent or a specific sight-reading talent. Rather, they are the results of deliberate long-term involvement in relevant domain-related activities and appropriate self-imposed challenges" (p. 25).

Meinz & Hambrick, 2010	
Type	Correlational
Test Construction	Author-composed, and Hambrick & Oswald (2005)
Production Method	Piano
Participants	57, average age was 30.9. 47.4% of subjects held one or more music degrees, with 80.7% considering piano their primary instrument.
Stimuli	Sight-reading: 6 pieces (*Four Star Sight Reading and Ear Tests* series (Berlin & Markow, 2002), WMC: Text
Pre- and Post-Tests	No
Control Group	No
Scoring System	Expert, 2 experienced and educated judges

Meinz & Hambrick (2010) evaluated sight reading ability with respect to working memory capacity, particularly after deliberate practice. For the sight reading portion of the study, the authors chose six musical excerpts from a sight reading method book at three levels of difficulty. One piece from each level indicated technical ability, and the other indicated the capacity for expression. The adult subjects, who were all practiced pianists, performed the piece at a given tempo after viewing it for 60 seconds, and then attempted the piece a second time after another 60 seconds of review. The sight reading task was recorded for each subject and scored independently by "two expert raters, each of whom taught piano at the university level and held a graduate degree in music" (p. 916). Each excerpt was scored on a scale of one to seven (though a zero score was possible if the piece was too difficult for a subject) by three measures: proficiency, musicality, and overall performance. With regard to working memory capacity, the authors did not find any correlation with sight reading nor with deliberate practice. However, they presume that working memory capacity has a part in sight reading ability "by determining the extent to which pianists can prepare for future keystrokes by looking ahead in music scores" (p. 917).

Woodwind and Brass Instrument Production

Elliott, 1982a	
Type	Correlational
Test Construction	WFPS (Watkins & Farnum, 1954), Author-composed technical proficiency evaluation, Criterion sightsinging test (Thostenson, 1970), Rhythm reading test (Boyle, 1968)
Production Method	Instrumental
Participants	30, undergraduate music theory students
Stimuli	WFPS Form A, and other material according to their tests
Pre- and Post-Tests	No
Control Group	No
Scoring System	Expert, 2 examiners (WFPS and Criterion test), 3 examiners (technical proficiency). Examiners were graduate music theory or education students.

To study seven areas that predict instrumental sight reading achievement, Elliott (1982a) employed the WFPS (Watkins & Farnum, 1954), the Criterion Sightsinging Test (Thostenson, 1970), the Rhythm Reading Test (Boyle, 1968), and a technical proficiency test. Elliott examined seven variables: technical proficiency, sight-singing skill, rhythm reading skill, and cumulative GPA scores for music theory, performance juries, and the subject's primary instrument, as well as total cumulative GPA. A trained graduate student administered all tests. Two graduate students studying instrumental education scored the WFPS and the Rhythm Reading Test. The WFPS was scored according to its manual. The Rhythm Reading Test was scored according to Boyle (1968), with examiners noting measures containing an error and totalling the correctly-performed measures. The Criterion Sightsinging Test was scored by two graduate students who taught sight-singing, and the *rhythm phrases* portion was scored for rhythm, while the *melodic phrases* portion was scored for pitch and rhythm. Three instrumental education graduate students scored the technical proficiency test, which consisted of scales, arpeggios, and patterns of thirds for performance, in every major key. For scoring, each task was rated on a scale of zero-to-seven, and measured for proficient technique and consistent tempo. Elliott found a strong correlation between subjects' rhythm-reading ability and sight reading ability, and he concluded that juried performance scores with rhythm-reading skill are the best predictors for sight reading achievement. However, Elliott writes, "it may be that those two tests [that is, the rhythm-reading and technical proficiency tests] were measuring a third skill (perhaps some sort of a pattern detection skill) that could be a determining factor in sight-reading ability" (p. 14). He suggests that more research in this area is needed.

Thompson, 1985	
Type	Correlational
Test Construction	WFPS (Watkins & Farnum, 1954) and author-composed
Production Method	Instrumental
Participants	30, ages 17-31 from community and university psychology classes
Stimuli	WFPS Form A, randomly-generated music in 24 measures, 7 selections from Gekeler (1969) and Voxman (1954), nonmusical letters
Pre- and Post-Tests	No
Control Group	No
Scoring System	Expert, author-scored

Thompson (1985) used the Watkins-Farnum Performance Scale and a composed piece to measure sight reading ability and the average duration of time between read notes, respectively. Subjects also completed a musical background questionnaire and additional tasks to measure recall, music- and letter-encoding, and choice reaction time. For the sight reading portion of the study, subjects first performed Form A of the WFPS; the exercises had been divided to be played at two different times throughout the experiment. It is assumed that the WFPS was scored according to its manual. The author recorded the length of time required to play the exercises, and divided the time by the number of measures. Thompson describes the additional calculations: "In order to take both speed and accuracy into account, an index of sight-reading skill was computed by dividing the number of correctly played measures...by the total number of seconds required to play all measures. This quantity, the number of correctly played measure per second (CMS) was used as the index of sight-reading ability" (p. 11). The second sight reading task, which comprised 24 measures of arbitrarily placed notes ranging from C-A, was calculated for the average time between notes. This data was compared with the choice reaction task, where a single note was presented in a series of trials and the subjects were to perform the note as quickly and accurately as possible. Here, the reaction times were recorded. Thompson found that the following factors correlated with sight reading ability: eye-performance span, music-encoding, and choice reaction time with individual notes.

Miller, 1988	
Type	Correlational
Test Construction	(MAT 2)(Colwell, 1986), Drake Musical Aptitude Tests (Drake, 1954), Measure of Musical Abilities (Bentley, 1966), Musical Background Test (Svenglais, 1978, unpublished), Music Attitude Inventory (Tomcala, 1977, unpublished), Group Embedded Figures Test (Oltman, Raskin, & Witkin, 1971), Frostig Movement Skills Test Battery (Orpet, 1972), Author-composed Music Student Questionnaire, WFPS (Watkins & Farnum, 1954), Author-composed Cognitive Test
Production Method	Instrumental
Participants	123, late-elementary students
Stimuli	Material according to the specific tests, as well as elementary band repertoire
Pre- and Post-Tests	No
Control Group	No
Scoring System	Expert, author-scored, with 2nd judge for reliability

Miller (1988) measured sight reading and rehearsed reading of wind instrumentalists aged 11 to 13 years of age, with the focus on pitch, rhythm, expression, and in combination. He examined musical abilities and individual features that were related to indicators of music reading skill, including grasp of tonality, pitch and rhythm audiation, the ability to continue a steady beat, musical memory, musical background, musical approach, field dependence and independence, hand-eye coordination, and musical aptitude. Miller employed a specific test to measure each skill, as well as the WFPS (Watkins & Farnum, 1954) to measure sight reading ability, though both the administration and scoring of the test were modified. Miller defines rehearsed reading as "the performance from notation of unmemorized music after sight-reading, practicing, and/or imagining the sound of the music from notation" (p. 5). In order to have consistency among subjects, there was a given amount of time to rehearse. Selected exercises from the WFPS, chosen for their broad range of difficulty, were given: five for the sight reading task and two of these for the rehearsed reading task. Markings of dynamics and expression were added to all five exercises. For scoring, Miller used an alternative method provided in the WFPS manual, where errors of pitch, rhythm and expression are recorded independently for each measure. Therefore, the correct measures are recorded, which Miller labeled the 'normal' score. Both reading tasks were scored in the same manner. Though Miller scored these unaided, a second evaluator scored 10% of the sight reading tests to confirm scoring accuracy, with inter-judge correlations between $r = 0.97$ and $r = 0.99$ for the four scores (that is, pitch, rhythm, expression, and 'normal').

Miller found that rhythmic aural-visual discrimination was more associated to instrumental sight reading and rehearsed reading ability than with melodic discrimination, and the cognitive music skills increase as the difficulty of repertoire increases. He notes that a subject's individual features are predictors of reading success, and that a cognitive model includes three groups of abilities: cognitive music skills, aural-visual rhythmic skills, and rhythm-pattern comprehension.

McPherson, 1994; 1995, McPherson, Bailey, & Sinclair, 1997	
Type	Correlational
Test Construction	WFPS (Watkins & Farnum, 1954), Australian Music Examinations Board (AMEB, 1990)
Production Method	Instrumental
Participants	101, secondary students
Stimuli	WFPS Material
Pre- and Post-Tests	No
Control Group	No
Scoring System	Expert, author-scored

McPherson (1994; 1995) and McPherson, Bailey, and Sinclair (1997) published findings from or related to an administration of the WFPS with clarinet and trumpet performers in high school. The objective of the 1994 study was to identify factors influencing sight reading ability. McPherson (1995) identified five performance skills: sight reading, performing rehearsed music, playing from memory, playing by ear, and improvising. To define the amount of correlation between these and 16 factors associated with musical background, subjects completed a questionnaire regarding the factors, and participated in tests evaluating sight reading, the ability to play from memory, the ability to play by ear, and the ability to improvise. The scores from the memory, ear, and improvisation tests were compared with the WFPS data. As well, McPherson studied 16 musical experience variables in relation to the performance skills measured by the tests: sight reading, rehearsed reading, playing from memory, playing by ear, and improvising. In the 1997 study, the postulated associations between the variables and performance skills were tested and further examined. Sight reading ability was measured with Form A of the WFPS. A pilot study (McPherson, 1994) provided inter-judge reliability and training and practice in the scoring system. McPherson scored each performance twice, and an “experienced independent musician” (p. 222) also scored each performance. McPherson’s two scores had a reliability coefficient of r = 0.99. The inter-judge reliability was found to be r = 0.98. Following the pilot study, McPherson alone recorded and scored the performances of the main experiment. McPherson describes the scoring method: “consistent with a procedure used by Elliott (1982b) and described in the WFPS scoring procedure as an alternative to drawing a cross through each incorrect measure, errors were noted using an abbreviated system of symbols above each measure in which an error or errors occurred” (p. 222). McPherson found that there were generally a larger number of rhythm errors, and that skilled sight readers were aware of score details such as time and key signatures, and aware of difficult portions of the piece. McPherson noted that sight reading and rehearsed reading were two separate forms of reading, and should be treated as such, particularly with novice musicians: “ability to sightread should not be considered synonymous with the student’s overall ability to perform a repertoire of rehearsed music” (p. 228).

McPherson (1995) gives more detail regarding the scoring of rehearsed music. It states that “specialist examiners” (p. 145) used a 12-point scale to assess the performances, with ‘1’ as the lowest score. This study also found a greater correlation between sight reading ability and improvisation ability than with sight reading and the performance of rehearsed music. McPherson remarks that this finding is notable, as the majority of subjects were not experienced in improvisation. For McPherson, Bailey, and Sinclair (1997), further study of the variables influencing performance skills showed that rehearsed music performance is directly affected by sight reading skill, as well as length of study, the ability to play from memory, and the ability to play by ear. McPherson and colleagues also found a "strong path coefficient" (p. 115) between the ability to sight read and the ability to play by ear. The authors write that this finding "is particularly noteworthy, and it is consistent with the hypothesis that the ability to play by ear exerts an important influence on the ability to sight-read" (p. 115-116). The ability to play by memory was also found to have an association with sight reading ability.

McPherson, Bailey, and Sinclair (1997) furthered the research of McPherson (1995) by grouping the 16 musical background factors into just four factors: early exposure, enriching activities, length of study, and quality of study, and undertook path analyses of the factors. However, the path between early exposure and sight reading was removed from the study due to "negligible path coefficients" (p. 113). With regard to sight reading ability, the authors found significant path coefficients between this skill and rehearsed performance, playing by ear, and playing from memory. Length of study also influences sight reading ability. Kopiez and Lee (2006) note that the studies by McPherson (1994; 1995) and McPherson and colleagues (1997) confirm that correlations between the performance skills are positive. Therefore, proficiency in music reading also means proficiency in these performance areas.

Eastlund Gromko, 2004	
Type	Correlational
Test Construction	WFPS (Watkins & Farnum, 1954), AMMA (Gordon, 1989), Schematizing Test (Holzman, 1954), Kit of Factor-Referenced Cognitive Test (Ekstrom et al., 1976), Iowa Tests of Educational Development (Hoover et al., 2003)
Production Method	Instrumental
Participants	98, secondary students
Stimuli	Material according to the specific test
Pre- and Post-Tests	No
Control Group	No
Scoring System	Expert, unknown

Eastlund Gromko (2004) examined the relationships between sight reading and other musical and nonmusical abilities. Musical abilities included tonal and rhythmic audiation, visual field articulation, and spatial orientation and visualization, and a series of specific published tests measured these abilities. Each factor was measured by a series of tests, including the Advanced Measures of Music Audiation (Gordon, 1989) and the WFPS (Watkins & Farnum, 1954). The WFPS assessed the sight reading skill of 98 high school wind instrumentalists. It does not specify who administered or scored the WFPS, but students were tested individually outside of their class times, so it is assumed that the primary investigator completed this portion of the study. As all other tests in this study were evaluated according to their procedures, it is also assumed that the WFPS was administered and scored according to its test manual. Eastlund Gromko found that the ability to sight read can be projected by a composite of cognitive skills, including rhythmic sight reading, audiation with aural feedback, spatial orientation, and the comprehension of text reading.

Hayward & Eastlund Gromko, 2009	
Type	Correlational
Test Construction	WFPS (Watkins & Farnum, 1954), author-adaptation of Elliott (1982), AMMA (Gordon, 1989), 3 subtests from Kit of Factor-Referenced Cognitive Tests (Ekstrom, French, Harman, & Dermen, 1976)
Production Method	Instrumental
Participants	70, university students (with 62 of the participants being music students)
Stimuli	Material according to the specific test
Pre- and Post-Tests	No
Control Group	No
Scoring System	Expert, 1 graduate student (WFPS)

Hayward and Eastlund Gromko (2009) studied predictors of music reading achievement: technical skill, aural discrimination, and spatial-temporal reasoning, with the objective of determining the degree of each factor. The factors necessitated kinesthetic/aural processing, aural/spatial processing, or visual/spatial processing, respectively. To determine the sight reading ability of the 70 instrumentalist subjects, the authors administered the WFPS. Subjects also completed an adaptation of Elliott's (1982a) technical proficiency test, The Advanced Measures of Audiation (Gordon, 1989), and subtests from the Kit of Factor-Referenced Cognitive Tests (Ekstrom, French, Harman, & Dermen, 1976). For the WFPS, subjects were tested individually and the test was scored during performance for errors of pitch and rhythm by "a graduate student with experience as a band director, test administrator, and wind band performer" (p. 30). The authors found that the three factors operated in combination to predict, "speed and accuracy of music sight-reading" (p. 33), with aural discrimination and spatial-temporal reasoning working together and technical proficiency working independently. They write, "although aural-spatial skills and technical proficiency skills are orthogonal or separate, they are both essential to the complex task of sight-reading" (p. 33). Hayward and Eastlund Gromko propose that this demonstrates the fact that reading and playing are components of sight reading.

Rhythm Reading Production

Levy, 2001	
Type	Correlational
Test Construction	Author-composed
Production Method	Instrumental rhythm reading
Participants	36 (main); 11 (for miscue analysis), all Grade 7 students in second year of instrument study
Stimuli	Standard band repertoire at beginning and intermediate levels
Pre- and Post-Tests	No
Control Group	No
Scoring System	Expert, by the author and 2 others

Levy (2001) states the purpose of her study was to examine the relationship between rhythmic structure and music reading. Levy proposed a music cognition model, or the Rhythm Reading Interactive Model (RRIM), adapted from language reading. "A reader's knowledge interacts with cues in music notation and its performance to effect rhythm reading. Aural and visual cue systems are readers' methods for structuring rhythm according to figural (functional, relational) and metric (beat-related, proportional) relationships" (p. v). Using a series of etudes from standard band repertoire, Levy assessed the sight reading ability of 36 late-elementary wind or brass instrumentalists. The subjects reported on their musical experience and training, participated in a reading interview concerning their perception and ideas about rhythm, completed a *Notation Practices and Training Questionnaire*, which comprised details about music reading and other musical activities, and completed two of Colwell's Music Achievement Tests. Levy also observed the subjects during their band rehearsals. Finally, each subject completed the sight reading task. For the sight reading test, Levy was concerned with rhythm reading miscues and writes, "to reduce the likelihood that pitch-, range- and articulation-related performance demands could obscure participants' rhythm reading process, I matched etudes' technical content with the learning sequence of the band method books used in participants' band classes" (p. 102).

Levy chose a sample of 11 of the recorded sight reading performances for additional rhythm reading miscue analysis by the experimenter and two judges, described as "educated and experienced instrumental music teachers" (p. 119). Levy then compared the two transcriptions and a third judge arbitrated the disagreements, which are described as "instances in which a performance of the two judges' notations would produce different rhythms, pauses, and tempo ebbs or flows" (p. 122). The scoring process for the rhythm-reading test began with miscue marking, with independent scoring of errors of pitch, rhythm, and articulation. *The miscue categories were* placed in a rubric and comprised substitutions, omissions, insertions, repetitions, pauses, and deviations from the tempo. The second portion of the scoring process was for miscue analysis, where Levy coded the miscues "within the context of the reader's musical phrase...in order to evaluate coherence and boundary assignment or phrase-making strategies" (p. 126).

Levy found that many areas were related to the use of cues, strategy, musical training, and musical ability, including skill in rhythm reading, approach to reading, and the collected performance data. Less-skilled readers were not as sensitive to aural cues and did not have a clear understanding of musical concepts. In contrast, skilled readers were sensitive to metric and aural cues and could consider the rhythmic component of music reading. Levy also found inventive and conventional reading approaches. Inventive readers used figurative pattern cues, while conventional readers used a broad range of cues.

3.2.4. Test Validation Studies

Rhythm Reading Production

Gutsch, 1966	
Type	Test Validation
Test Construction	Author-composed
Production Method	Rhythm Reading; Verbal Response
Participants	771, Grade 5 or older with at least 6 weeks instrumental experience
Stimuli	1 or more measures of rhythm patterns
Pre- and Post-Tests	Yes, for select participants
Control Group	No
Scoring System	Expert, unknown

Gutsch (1966) developed an assessment tool to measure rhythm reading. In this early study, the measurement and scoring details are not described. However, the test was designed for individual administration, and its purpose was to better understand instrumental music reading ability. A second purpose of the test was to differentiate between subjects, particularly with experience and age. The stimuli consisted of two forms of 100 flash cards each, containing “rhythmic problems” (p. 377). Each flash card presented one or two measures of rhythm. Each subject viewed and responded to one form of the test, followed by the second form of the test within 48 hours, though the mode of response is not specifically named. All performances were recorded. Gutsch administered this test to 771 middle- to late-elementary students, who had at least six weeks of instrumental training. After a year, 120 additional subjects participated in the test, with 81 of the initial subjects re-taking the test. Gutsch reports that the test was reliable and had construct validity, and that there was a high degree of scorer reliability ($r = 0.99$). He also found that experience had the greatest influence on achievement, with little influence from the age of the subject.

Sight-Singing Production

Thostenson, 1967	
Type	Test Validation
Test Construction	Author-composed
Production Method	Sight-singing
Participants	Unknown, but 150 +, undergraduate and graduate music students
Stimuli	Author-composed sight-singing and pitch, rhythm, melodic dictation tests
Pre- and Post-Tests	Yes
Control Group	No
Scoring System	Expert, 4 people for administration and evaluation

Thostenson (1967) administered the CSS76 Criterion Sightsinging Test and the PRM78, Pitch, Rhythm, Melodic Dictation test in a pilot study to 54 university students. The author-constructed tests are for music students at the undergraduate and graduate level (Roach, 1982) and the objectives are to observe increased proficiency with music dictation with ear training, and to correlate proficient dictation with sight singing ability (Thostenson, 1967). The first area of measure in the CSS76 Criterion Sightsinging Test is concerning pitch phrases, where a two-note interval or four-note melodic phrase is presented and performed. For the measure of rhythm phrases, there are melodic segments of two measures without pitch information. For the measure of melodic phrases, both pitch and rhythmic information is present in a three-to-nine-note phrase, which also includes a time signature and tempo markings, and uses a sung syllable for performance (Elliott, 1982a). All phrases are written in the treble clef and span a range of two octaves (Thostenson, 1967). The CSS76 Criterion Sightsinging Test was employed in a series of tests administered by Elliott (1982a). Elliott describes the scoring method: for pitch and rhythm phrases, items are evaluated on the correct or incorrect pitch or rhythm only, respectively. "For the 'melodic phrases' subjects are given both a starting pitch and a starting tempo and items are scored on the basis of both pitch and rhythm" (p. 8). The subjects may also choose to repeat each phrase, and the test administrator accordingly labels each score as *first* or *second*. The correct responses are totaled, with a composite score obtained from the three measures. The Criterion Sightsinging Test offers a high degree of reliability.

Thostenson (1967) found that the test was reliable, with an averaged inter-correlation coefficient of r = 0.86. The main study sought to correlate the full PRM78

Dictation Test, as well as subdivisions of the test, with sight-singing proficiency as measured by the CSS78 Criterion Test. Another objective of the study was to identify the average ability of students for both dictation and sight-singing. Thostenson collected data of the two tests from a large number of subjects, both at the undergraduate and graduate levels, at various universities. The PRM78 was administered early in the school year, with a second administration at the end of the year. The CSS78 was administered only at the end of the year. Thostenson found that graduate students were superior with pitch components and the sight-singing test in general, while undergraduate subjects had the lowest scores with the rhythm components. He also found that subjects studying keyboard instruments had slightly increased scores. Thostenson noted that assessing rhythm is more difficult than pitch.

Henry, 2001	
Type	Test Validation
Test Construction	Author-composed
Production Method	Sight-singing
Participants	42 (pilot); 183 (main), all secondary students
Stimuli	Author-composed tonal melodies containing pitch skill patterns
Pre- and Post-Tests	No
Control Group	No
Scoring System	Expert, author, with a second scorer for reliability

Henry (2001) developed and administered the Vocal Sight Reading Inventory (VSRI) test to 42 high school students for the pilot study, and 183 high school students for the main study.

The scale assesses vocal sight-singing ability, with additional objectives of allowing "formative and summative evaluation" (p. 21) and establishing "validity, reliability, and normative data" (p. 21). The VSRI is designed for individual administration for high school students; as such, another objective was that the duration of the test be compatible in the choral classroom setting. There are two forms of the test and each consists of eight major key melodies notated in the bass and treble clefs and without complex rhythms. The melodies contain embedded pitch skills, which Henry obtained from tonal music. She named seven categories in which to classify the pitch skills, or functional melodic patterns: conjunct, which are simple ascending or descending patterns; tonic; dominant; subdominant; cadential, with four authentic cadences; modulatory; and chromatic. The test itself requires subjects to listen to a chord progression of the key and the first pitch of the melody. They view and, if desired, practice the melody for 30 seconds before the chord progression and first pitch are played again, and then sing the melody. The scoring procedure applies to the pitch patterns within the melody, by codes associated with each pitch. Incorrect pitches are scored, though the final score given to the subject represents the number of correct component skills executed. For the 2001 study, Henry and a second scorer evaluated recordings of the pilot performances with the scoring method proposed by the VSRI. For the main study, Henry scored the VSRI with the proposed method and a second scorer independently scored using a note-by-note method, for inter-scorer reliability. This was found to be r = 0.97. Henry (2004) has since administered the VSRI, and Killian and Henry (2005a) administered a modified version of the test.

During the 2001 study, for each performance subjects were given instruction and the melodies, and after the key was established, they studied the stimuli for 30 seconds before singing the entire melody. Henry found consistently low scores for the VSRI. Henry posits that this could be due to the large number of subjects participating in the study, or for unknown reasons. However, the VSRI was found to be valid and reliable as an assessment tool.

3.3. Summary

For **cognitive process studies**, errors are more common with less-skilled readers (Salis, 1977; 1980). Elliott (1982b) found that errors of rhythm were more common for this subject group, while Gudmundsdottir (2003; 2010a) noted that pitch errors were common for this subject group, and particularly in the left hand for keyboardists. Experience allows for continued performance, even when errors are committed (Drake & Palmer, 2000). Visual and auditory feedback plays a role in reading and performing success, including continued physical movement (Drake & Palmer, 2000; Fine, Berry, & Rosner, 2006; Havlicek, 1968). Music reading achievement is aided by technical proficiency (Gilman, 2000), pattern recognition ability (Fine, Berry, & Rosner, 2006), sight reading experience and reaction time (Lee, 2004), mental speed (Kopiez & Lee, 2008; Kopiez et al., 2008), and flexibility of achievement factors as the task requirements change (Kopiez & Lee, 2006). Music literacy consists of production, perception and reflection factors; independent of each other (Scripp, 1995), and literacy is achieved in keyboard musicians with an instinctive vertical-to-horizontal stimulus response (Stewart, Walsh, & Frith, 2004).

For **instructional studies**, success in music reading is achieved by training in pattern recognition (Pike & Carter, 2010; MacKnight, 1975; Zhukov, 2014a; Henry, 2004), pattern training with vocalization, and aural and visual perception skills (MacKnight, 1975; Grutzmacher, 1985; 1987). Instruction in rhythm reading is particularly important to music reading success (Boyle, 1968; Zhukov, 2014a; Palmer, 1976; Bebeau; 1982; Salzberg & Wang, 1989). General music education benefits music reading success (Micheletti, 1980; Kostka, 2000).

For **correlational studies**, music reading achievement can be predicted by rhythm reading skill (Elliott, 1982a; Miller, 1988; Levy, 2001), music-encoding skill and reaction time (Thompson, 1985), aural skill (Kornicke, 1992; McPherson, 1995; McPherson, Bailey, & Sinclair, 1997; Eastlund Gromko, 2004; Hayward & Eastlund Gromko, 2009; Erlings, 1977), psychomotor skill (Eaton, 1978), specialization and experience (Lehmann & Ericsson, 1993; 1996; Thostenson, 1967; Gutsch, 1966); and deliberate practice (Meinz & Hambrick, 2010).

Table 9. *Synthesis of summary tables for cognitive process studies.*

	Test Construction	Production Method	Participants	Pretest/ Posttest	Control Group
Salis 1977; 1980	Author-composed	Piano	26 (adults)	No	No
Banton 1995	Author-composed	Piano	15 (proficient)	No	No
Drake & Palmer 2000	Author-composed	Piano	60 (children, adults)	No	No
Gilman 2000	Author-composed	Piano	24 (proficient)	No	No
Gudmundsdottir 2003; 2010a	Author-composed	Piano	35 (children)	No	No
Stewart, Walsh, & Frith 2004	Author-composed	Computer (acting as piano)	26; 22 (university)	No	No
Lee 2004; Kopiez & Lee 2006; 2008; Kopiez et al. 2006	Author-composed	Piano	52 (university)	No	No
Havlicek 1968	Author-composed	Piano; Instrumental	21 (university)	No	No
Elliott 1982b	WFPS	Instrumental	30 (university)	No	No
Shehan 1987	Author-composed	Percussion; Rhythm reading	49 (children)	No	No
Halpern & Bower 1982	Author-composed	Transcription	24; 12; 14 (proficient/not proficient)	No	No
Pick, Unze, Metz, & Richardson 1982	Author-composed	Repetition	60; 44 (children)	No	No
Scripp 1995	Author-composed (adapt WFPS)	Sight-singing	20 (university)	Longitudinal	No
Fine, Berry, & Rosner 2006	Author-composed	Sight-singing	22 (proficient)	No	No

Table 10. *Synthesis of summary tables for instructional studies.*

	Test Construction	Production Method	Participants	Pretest/ Posttest	Control Group	Duration
Lowder 1973	Author-composed	Piano	23 (university)	No	Yes	1 semester
Micheletti 1980	Author-composed; Shake Piano Test	Piano	40 (university)	Yes	Yes	10 weeks
Betts & Cassidy 2000	Author-composed	Piano	39 (university)	Yes	No	1 semester
Kostka 2000	Author-composed	Piano	69 (university)	Yes	Yes	10 weeks
Pike & Carter 2010	Adapt WFPS	Piano	43 (university)	Yes	Yes	12 weeks
Zhukov 2014a	Author-composed	Piano	100 (proficient)	Yes	Yes	10 weeks
MacKnight 1975	WFPS; MAP; MAT; Author-composed	Instrumental	Approx. 90 (children)	Yes	Yes	32 weeks
Grutzmacher 1985; 1987	Author-composed; Adapt WFPS; MAP; ITML	Instrumental	48 (children)	Yes	Yes	14 weeks
Salzberg & Wang 1989	Author-composed	Instrumental Rhythm reading	46 (children)	Yes	No	Unk.
Boyle 1968	Author-composed; Adapt WFPS	Instrumental Rhythm reading	201; 191 (secondary)	Yes	Yes	14 weeks
Palmer 1976	Rhythm methods; MAP; MAT; ITML	Rhythm reading	136 (children)	Yes	Yes	5 months
Bebeau 1982	Author-composed	Rhythm reading	27; 80 (children)	Yes	No	18 days
Demorest 1998	Author-composed	Sight-singing	306 (secondary)	Yes	Yes	1 semester
Henry 2004	VSRI	Sight-singing	67 (secondary)	Yes	No	12 weeks

Table 11. *Synthesis of summary tables for correlational studies.*

	Test Construction	Production Method	Participants	Pretest/ Posttest	Control Group
Erlings 1977	Author-composed; MAP; MAT	Piano	33 (university)	Yes	No
Eaton 1978	Author-composed	Piano	73 (university; adults)	No	No
Kornicke 1992	Author-composed; Nonmusic tests	Piano	73 (university; adults)	No	No
Lehmann & Ericsson 1993	Author-composed	Piano	16 (university)	No	No
Lehmann & Ericsson 1996	Author-composed	Piano	16 (university)	No	No
Meinz & Hambrick 2010	Author-composed	Piano	57 (proficient)	No	No
Elliott 1982a	WFPS; Author-composed; Criterion; Boyle Rhythm	Instrumental	30 (university)	No	No
Thompson 1985	WFPS; Author-composed	Instrumental	30 (university; adults)	No	No
Miller 1988	MAT; Drake; Bentley; WFPS; Etc.	Instrumental	123 (children)	No	No
McPherson 1994; 1995; McPherson et al. 1997	WFPS; AMEB	Instrumental	101 (secondary)	No	No
Eastlund Gromko 2004	WFPS; AMMA; Etc.	Instrumental	98 (secondary)	No	No
Hayward & Eastlund Gromko 2009	WFPS; Adapt Elliott; AMMA; Etc.	Instrumental	70 (university)	No	No
Levy 2001	Author-composed	Instrumental Rhythm reading	36; 11 (children)	No	No

Table 12. *Synthesis of summary tables for test validation studies.*

	Test Construction	**Production Method**	**Participants**	**Pretest/ Posttest**	**Control Group**
Gutsch 1966	Author-composed	Rhythm reading	771 (children)	Yes, for select	No
Thostenson 1967	Author-composed	Sight-singing	150 + (university)	Yes	No
Henry 2001	Author-composed	Sight-singing	42; 183 (secondary)	No	No

Table 13. *Summary of number of studies by classification.*

Measurement	**Cognitive Process**	**Instructional**	**Correlational**	**Test Validation**
Number of Studies	**17**	**14**	**18**	**3**

CHAPTER 4

4. Eye-Tracking Measurements

According to Kopiez and Galley (2002), there has been thorough study of "visual information processing" (p. 683). The study of eye movement with eye tracking is necessary for the biological and medical understanding of the oculomotor system, and is increasing in many fields, including psychology and education, due to its novel data collection methods and ability to encourage discussion about vision and experience (Reulen et al., 1988; Horsley et al., 2014). Eye movement provides understanding for human behaviour (Reulen et al., 1988), and Tatler and colleagues (2014) offer support for this: "first, the locations selected for fixation provide us with insights into the changing moment-to-moment information requirements for the behaviours we engage in. Second...eye movements provide an ideal and powerful objective measure of ongoing cognitive processes and information requirements during behaviour" (p. 3). In the field of psychology, where music reading studies are found, eye-tracking research provides evidence of the connection between eye-movement and the cognitive processes involved in reading (Gilman, 2000; Madell & Hébert, 2008). Specifically, eye-tracking studies measure eye-movements while reading musical notation: eye-movement behavior under various conditions (Salis, 1977; Sloboda, 1977; Goolsby, 1994a; Kinsler & Carpenter, 1995; Gilman, 2000; Kopiez & Galley, 2002; Drai-Zerbib & Baccino, 2005; Drai-Zerbib, Baccino, & Bigand, 2012; Penttinen & Huovinen, 2011; Penttinen, Huovinen, & Ylitalo, 2012), parafoveal vision (Gilman, 2000), perceptual span (Goolsby, 1994a; Truitt, Clifton, Pollatsek, & Rayner, 1997; Waters, Underwood, & Townsend, 1998), eye-hand span (Sloboda, 1974; Thompson, 1985; Truitt, Clifton, Pollatsek, & Rayner, 1997), pattern recognition (Young, 1971; Waters, Underwood, & Townsend, 1998), and comparison of reading skill in relation to visual processing among subjects (Bean, 1938; Young, 1971; Sloboda, 1974; 1976a; Salis, 1977; Goolsby, 1994a; 1994b; Truitt et al., 1997; Waters, Underwood, & Findlay, 1997; Waters & Underwood, 1998; Furneaux & Land, 1999; Gilman, 2000; Kopiez & Galley, 2002; Gilman & Underwood, 2003; Drai-Zerbib & Baccino, 2005; Burman & Booth, 2009; Penttinen & Huovinen, 2011).

Eye-tracking apparatus is regularly used with eye movement research. According to Reulen and colleagues (1988), "interest in oculomotor research has confronted the recording systems with numerous requirements, such as mechanical and electronic stability, easy and non-traumatic application, non-contacting to the eye, low cost, ample linear properties, and sufficiently wide dynamic measuring range and high spatial resolution" (p. 20). From the available eye movement literature relating to music reading, there are three categories in which to classify studies by apparatus used: 1. Viewing devices (Bean, 1938; Pagan, 1970; Young, 1971; Sloboda, 1976a; 1978b; Salis, 1977, Burman & Booth, 2009; D'Anselmo, Guiliani, Marzoli, Tommasi, & Brancucci, 2015; Sloboda, 1974; 1977; Thompson, 1985), 2. Eye-tracking devices (Goolsby, 1994a; 1994b; Truitt et al., 1997; Gilman, 2000; Penttinen & Huovinen, 2011; Penttinen, Huovinen, & Ylitalo, 2012; 2015; Drai-Zerbib, Baccino, & Bigand 2012; Drai-Zerbib & Baccino, 2005; Gilman & Underwood, 2003; Ahken, Comeau, Hébert, & Balasubramaniam, 2012; Kinsler & Carpenter, 1995; Waters, Underwood, & Findlay, 1997; Waters & Underwood, 1998; Furneaux & Land, 1999) and 3. Electrooculography (Kopiez & Galley, 2002; Schön & Besson, 2002; 2005; Gunter, Schmidt, & Besson, 2003; Wong, 2010). First, we discuss eye movement and specific measurement of eye movement during music reading research.

Eyes move in a succession of fixations and saccades while viewing a stimulus (Penttinen & Huovinen, 2011). White and Hall (2012) classify the five basic eye movements into those that modify gaze: "saccades, smooth pursuit movements, and vergence movements" (p. 438), and those that stabilize gaze: "vestibulo-ocular and optokinetic movements" (p. 438).

According to Gilman (2000), the fovea, a part of the retina with the clearest vision, requires eye movements; namely fixations, to collect the visual material in both text and music reading. The retina is the inmost of the eye's three layers, and is the only region in the eye with the ability to transfer signals from its light-sensitive neurons. During fixation, the eyes are almost motionless and thus acuity of the eye is ideal (Fitzpatrick & Mooney, 2012). Tatler (2014) notes that this area of acuity is a limited resource, and thus requires careful and suitable distribution as the environment demands. However, during fixation one is able to learn much about the stimulus such as its location, size, direction, and texture, among other features (Fitzpatrick & Mooney, 2012).

Many music reading studies support fixation length at 200-400 milliseconds, including Goolsby (1994a; 1994b), Truitt and colleagues (1997), Waters and colleagues (1997), and Waters and Underwood (1998). However, eye-tracking research has consistently found that increased reading skill results in shorter fixations (Goolsby, 1994a; Truitt et al., 1997; Waters & Underwood, 1998; Waters et al., 1998), and an expertise effect of fewer fixations has been identified (Drai-Zerbib & Baccino, 2005).

The significant areas of measure with eye-tracking studies are, "fixation durations and number of fixations as these measures are often reliable indicators of processing difficulty...gaze duration (total time spent looking at a particular target), and first fixation duration (the duration of the first fixation on a given target)" (Gilman, 2000, p. 24). Madell and Hébert (2008) remark that measuring fixation position is especially significant; in text reading, it is known that a word is fixated, but in music reading the principle fixation point or points is still undetermined. This may be of particular importance because musical notation, and thus its meaning, is more condensed on a score than are printed words on a page (Kinsler & Carpenter, 1995). Additionally, fixation duration corresponds to comprehension (Hyönä, 2011). Single areas of fixation are relatively correlated to features of the stimulus (Kinsler & Carpenter, 1995), while Truitt and colleagues (1997) suggest that musicians may use fixations to track upcoming notation.

Saccades support the fovea by moving the eye with quick motions to the next fixation, while upper motor neurons located in the superior colliculus, which is in the midbrain, control these movements. While the eye is fixating, other cells prevent the neurons from making unnecessary saccades (White & Hall, 2012). During saccade movement there is very little or no visual intake of information (Kinsler & Carpenter, 1995; Truitt et al., 1997; White & Hall, 2012). Saccades can move in any direction, and these movements are automatic when eyes are open, but can also occur willingly (White & Hall, 2012). Forward and backward saccades are called progressive and regressive, respectively, with 10-15% of saccades moving backwards (Gilman, 2000; Truitt et al., 1997).

Kopiez and Galley (2002) note that saccadic patterns are significant for the study of music reading and especially for music sight reading, and Gilman (2000) also identifies the spans of progressive and regressive saccades as items for measure. Saccades have been measured at 15-50 milliseconds and the duration is dependent on the length of the movement (Truitt et al., 1997). With reading, the timing of saccades is randomly determined and can change even with consecutive viewings of the same stimulus (Kinsler & Carpenter, 1995).

To preface three experiments exploring extrafoveal vision in her 2000 study, Gilman writes, "so far, relatively little attention has been paid to the nature of parafoveal processing during music reading" (p. 245). Fourie (2000) labels two key regions for the optical system: a central area with visual acuity and a peripheral area, which surrounds the central area and influences eye-motion in reading. Hyönä (2011) further partitions this outer area into the parafovea and periphery. "The parafoveal area extends up to 5° of visual angle to the right and left from the fixation point, while the visual periphery refers to the area beyond the parafovea...readers can extract useful information from the parafovea, whereas peripheral information is of very little use" (p. 819-820). Though, according to Gilman (2000), there is support that material in both the parafovea and periphery regions can be processed. It has now been established that a musical stimulus seen in parafoveal vision is processed to some extent before fixation (Gilman & Underwood, 2003).

The use of a 'moving-window paradigm' had been previously used in studies of text reading, as developed by McConkie and Rayner (1975), to measure extrafoveal vision; specifically by defining the smallest range for the window while allowing continuous reading performance (Gilman, 2000; Gilman & Underwood, 2003; Hyönä, 2011). This technique was then applied to music reading (Truitt, Clifton, Pollatsek, & Rayner, 1997; Gilman & Underwood, 2003). The moving-window technique is described by Gilman and Underwood (2003):

> During this paradigm, a stimulus is presented to be viewed by a participant, whose eye-movements are continuously recorded and updated. Within a predefined area ("window") around the point of fixation, all stimulus information is retained, while the information outside the window is occluded or degraded. The size and shape of the window is varied and the participant's performance is compared with performance in a control condition in which no window is presented (p. 202).

The moving window technique is used to measure the perceptual span. In regard to this technique, "the perceptual span is defined as the largest window for music preview that significantly distorts eye movement or performance parameters, delineated by the number of notes (or musical beats) displayed ahead of the current fixation" (Burman & Booth, 2009, p. 304). Similarly, Gilman and Underwood (2003) define it as "the minimum window size required in order for the participant's performance to remain uninhibited" (p. 202). A simpler definition is "the size of the effective visual field" (Truitt et al., 1997, p. 144) or "the span of effective stimulus" (Gilman & Underwood, 2003).

In regard to eye movement research, Madell and Hébert (2008) identify information regarding the perceptual span and eye-hand span as the most established research within music reading. They write, "this owes primarily to agreement amongst researchers as to what is being measured and why. Though the studies are not identical, they all agree on the fundamental definitions of the eye-hand span and perceptual spans" (p. 161). Penttinen and Huovinen (2011) note that all movement of the eye, regardless of what the eye is viewing, includes the perceptual span, but that eye-hand span "is especially related to the processing of musical notation during performance" (p. 198).

Madell and Hébert (2008) offer two similar definitions of the eye-hand span: the period or length between the production of music and perception of the notation, or between the production of music and the position of the eyes. According to Penttinen, Huovinen, and Ylitalo (2015), the eye-hand span can be "variously measured in time, number of beats, spatial distance, or number of notes" (p. 38). Studies have consistently proven that the eye-hand span increases with expertise (Truitt et al., 1997; Gilman & Underwood, 2003; Madell & Hébert, 2008). In agreement, Fourie (2004) considers a "wide eye-hand span" (p. 14) as advantageous for the music reader. However, the concept of the eye-hand span may not be easily taught to student pianists: "simply encouraging the student to 'look ahead' would not necessarily result in improvement in eye-hand span" (Zhukov, 2014b, p. 488).

The brain coordinates all of these eye movements as it receives electrical signals from the seeing eye and processes them in its retinal neural circuits (Fitzpatrick & Mooney, 2012). The information is processed in many areas of the brain; thus perception occurs while reflexes adjust the eyes to specific targets. This progression of events has been thoroughly studied, and scientists understand how "different classes of neurons within this pathway encode the variety of visual information...that we ultimately "see". The parallel processing of different categories of visual information continues in cortical pathways that extend beyond the primary visual cortex, supplying visual areas in the occipital, parietal, and temporal lobes. Visual areas in the temporal lobe are primarily involved in object recognition, whereas those in the parietal lobe are concerned with motion" (p. 257). However, the authors note that the foundational means to visual perception is not yet well known.

4.1. Eye Movement Method Descriptions

The eye-tracking apparatus is outlined by the viewing devices, various and infrared-based eye trackers, and electrooculography used in music reading assessment methods.

4.1.1. Viewing Devices

The Tachistoscope

Measurement of eye movements in music reading began with the use of the tachistoscope (Lowder, 1973). The tachistoscope was first introduced in 1859, but had already been utilized in physiology research prior to its introduction in literature (Benschop, 1998). Benschop explains that there have been many forms of the tachistoscope with varying manifestations; all associated by their use in "experimental practice" (p. 26). In the 1880s, a design modification allowed for greater accuracy in the function of the tachistoscope. Accordingly, it was used to measure reading rate (Coombs, 2012), initially for text reading studies and subsequently adopted for music reading studies (Salis, 1977), which includes the measure of eye movement in combination with an eye-tracking device.

Coombs (2012) writes that the tachistoscope is "a device designed to clock the speed of visual perception by flashing single images at an observer for fractions of a second" (p. 107), with the objective to isolate a distinct and separate visual object. The tachistoscope can have two or three fields for the presentation of stimuli (Sloboda, 1976a), and the images are displayed for a controlled amount of fixation time (Salis, 1977; Micheletti, 1980). When eye movement measure is required of the study, additional apparatus is needed to record data. Examples of eye movement apparatus used in conjunction with the tachistoscope include the use of an ophthalmograph, which photographs eyes during reading (Bean, 1938), a modified *Westgate Eye-Movement Camera*, which photographs light reflected off of the eye while simultaneously photographing what is being viewed (Young, 1971), the *Iscan*, an infrared eye-tracker mounted on the head to monitor the position of the eye (Burman & Booth, 2009), and a *Sony DCR-HC17E* video camera to inspect eye movements (D'Anselmo, Guiliani, Marzoli, Tommasi, & Brancucci, 2015).

Figure 3. *An example of a tachistoscope (Bean, 1938, p. 12).*

The Light-Out Technique

The 'light-out' technique can be considered a variant of the tachistoscope method as it also involves the visual presentation of stimuli. This technique was used in early music reading studies (Sloboda, 1974; 1977; Thompson, 1985) and employs a projector and multiple slides, which display the stimuli. The projector is turned off at a point unknown to the subject in order to obtain measurement. These early studies aided in establishing a foundation for measurement with the use of eye-tracking devices.

4.1.2. Eye Tracking Devices

Eye-tracking studies in the field of music reading research have used a variety of apparatus. Some are invented or adapted for their particular research objectives (Furneaux & Land, 1999). They may be altered from or modeled after previous use (Waters & Underwood, 1998). Devices that use infrared light are common, exemplified in the studies of Goolsby (1994a; 1994b), Kinsler and Carpenter (1995), Truitt and Colleagues (1997), Waters and Colleagues (1997), Waters and Underwood (1998), Gilman (2000), and Drai-Zerbib and Baccino (2005).

Various Eye Trackers

The Tobii 1750 Eye Tracker/Tobii TX300 Eye Tracker

Recent study of music reading has employed *Tobii Eye Trackers* (Penttinen & Huovinen, 2011; Penttinen, Huovinen, & Ylitalo, 2012; 2015; Drai-Zerbib, Baccino, & Bigand 2012). *Tobii* is a Swedish company founded in 2001, which has developed and continues to improve on eye-tracker technology. The eye-tracking equipment contains a sensor with the ability to locate precise eye fixation, as well as identifying one’s attention and awareness (www.tobii.com, retrieved 2016).

Figure 4. *The Tobii 1750 Eye-Tracker (http://pages.bangor.ac.uk/~pss034/ESRCProject/Eye-tracking.html).*

SR Eyelink

This video-based apparatus began as a head-mounted camera, and is now desk-mounted. It is comprised of three small cameras: two cameras track the dominant eye and the third follows the point of gaze. Eye movements are sampled at 500 Hz, with little noise. It is particularly suited to saccade and smooth pursuit measurements (sr-research.com, retrieved April 15, 2016). This device is common in higher education and has been used with music reading research (Gilman & Underwood, 2003; Ahken, Comeau, Hébert, & Balasubramaniam, 2012).

Land Recording System

Furneaux and Land (1999) used an eye-tracking device that had been designed and operated with an earlier study (Land, 1993). Furneaux and Land describe the apparatus, which is a recording system mounted on the head: "This non-intrusive system allows full and normal head and body movement during performance. A split-screen video recording was produced via two mirrors: a part-silvered mirror images the scene ahead, and a concave mirror images the subject's left eye" (p. 2436). Land (1993) explains that both the scene being viewed and eye movements are recorded simultaneously.

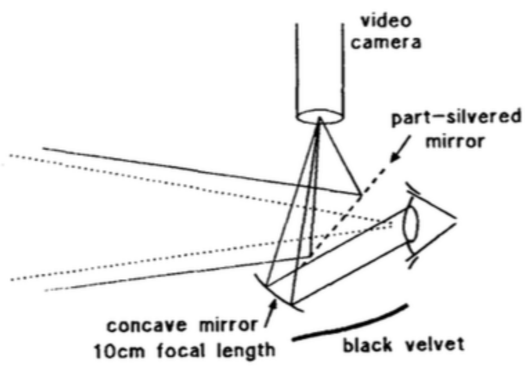


Figure 5. The Land recording system, showing how the apparatus captures eye movement (Land, 1993, p. 491).

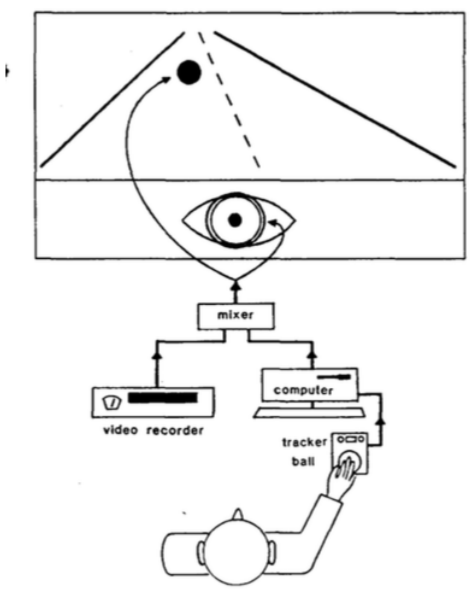


Figure 6. *The Land Recording System, showing how the apparatus is set up for a subject (Land, 1993, p. 491).*

Infrared-Based Eye Trackers

The Dual-Purkinje-Image Eyetracker

Development began on this instrument in 1965 at the Stanford Research Institute (SRI), and was licensed to a Fourward Technologies in 1988 for further improvement (http://www.w63309.wix.com). Gilman (2000) describes the eye-tracker as possessing an infrared light, which is beamed at the right eye and creates four reflections. These are called purkinje images, and the separation of the first and fourth images determines the angle of the eye. It possesses a two-dimensional field of view, which can identify the direction of the eye using infrared light to allow for natural vision (http://www.63309.wix.com). The *DPI Eyetracker* is described as, "a precise device with temporal sampling resolution of 1.0 ms" (Madell & Hébert, 2008, p. 160). Gilman (2000) notes that the device is extremely precise but that the subject's head must remain motionless during tracking. Therefore, nonverbal tasks are best for the eye-tracker and natural movement tasks, such as sight reading, are not feasible.

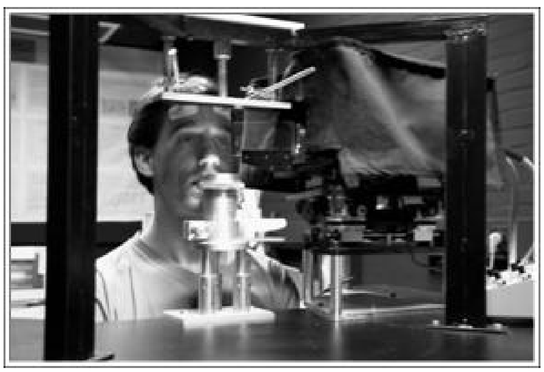


Figure 7. *The DPI Eye-Tracker, showing how the apparatus is set up for a subject (http://ppw.kuleuven.be/home/english/research/lep/resources/purkinje).*

Eye-Gaze Software

The apparatus has been used for music reading study (Drai-Zerbib & Baccino, 2005), and consists of a camera that reflects infrared light beamed onto the center of the cornea. The camera is placed below a computer monitor, which displays stimuli. It requires regular recalibration and measures eye movement at 60Hz, tracing the movement of the pupil in real time. *Eye-Gaze* determines the spatial coordinates of the eye from fixation and can then track and record movements. An advantage of this technology is that it allows natural physical head movement during measurement (www.eyegaze.com, retrieved March 29, 2016).

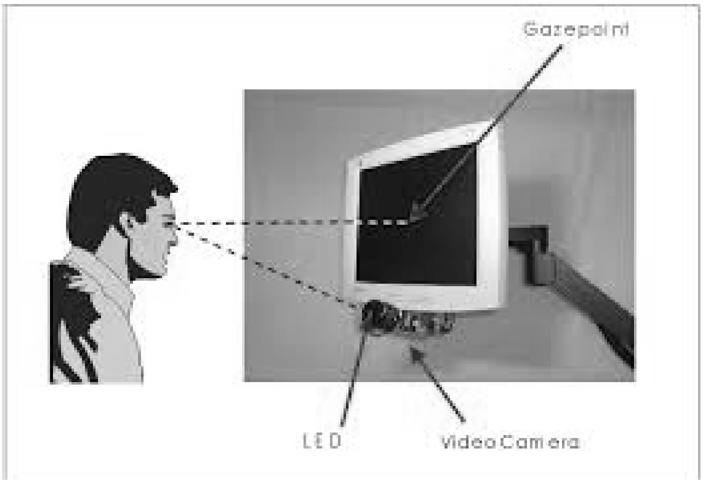


Figure 8. *EyeGaze Software, showing how the apparatus is set up for a subject (http://www.quadadapt.com/ins_specs.html).*

Carpenter Scleral Infrared Eye Movement Transducer

As described by Carpenter (1988), a scleral infrared eye movement transducer, which can be labeled a 'photoelectric viewer', views the border of the sclera and the iris of the eye, called the limbus. Simply, a point of light is beamed on the limbus while a photodetector collects the scattered light. However, Carpenter provides a history of improvements made to the device throughout the 1950s and 1960s: infrared light was used in place of radiation, and two points of light were used for recording and measurement instead of one. The lights were later projected in thin bars to differentiate from perpendicular eye rotations. An image of the eye was also illuminated and displayed on a screen. Finally, Rashbass (1960) employed an oscilloscope screen to trace the movement of the limbus. The scleral infrared transducer is effective to locate the position of the eye within 0.25 of a degree (Kinsler & Carpenter, 1995), and it can measure both the horizontal and vertical movements of the eye, including "a sensitivity of 3 inches of arc over some 30 degrees for horizontal movements, and 10 inches over 20 degrees for vertical movements (when the inner edge of the iris is used)" (Carpenter, 1988, p. 409). Rashbass (1960) lists the advantages of the device as its ability to adapt to varying measurement needs, including different numbers of subjects. It does not affect the eye or vision itself. Kinsler and Carpenter (1995) note that the subject's head must be secured with use of a bite-bar and the device must be calibrated regularly, as head and eye movements are not distinguishable from each other (Rashbass, 1960; Carpenter, 1988). Finally, the general experimental environment must have limited illumination.

Skalar IRIS System

Waters, Underwood and Findlay (1997) used an infrared eye-tracking system designed by Reulen and colleagues (1988). The *Skalar IRIS System* provides exact measurement for both horizontal and vertical movements and is thus arranged to trace the horizontal movement of one eye and the vertical movement of the other, with sampling every 5 milliseconds (Reulen et al., 1988; Waters et al., 1997). Infrared light is reflected from the limbus by way of nine infrared-light-emitting diodes with nine phototransistors, which correspond with the emitters. The diodes and sensitive detectors are located on the nasal and temporal sides of the eye, corresponding with the limbus, and are in a case within a light metal frame, to ensure that the vision is not obscured. Depending on the position of the case in the frame, horizontal or vertical measurement is possible (Reulen et al., 1988). According to Reulen, "the horizontal dynamic range of the IRIS eye-movement system is about 30 degrees lateral, measured from the centre position. Because of the anatomy of the eye and the eyelids, the IRIS measuring range for vertical eye movements is limited to about 25 degrees laterally" (p. 22). Eye position is determined by subtracting the signals from the nasal and temporal detectors, and this system provides consistent measurement of fixations, saccades, and smooth-pursuit and vergence movements. This device has a high resolution, and can measure eye movement in darkness and eye blinks. It emits little noise, does not have contact with the eye, and is easy to operate and align. Reulen notes that a limitation of the device is the possible interference of “natural pupillary oscillations” (p. 25).

Figure 9. *The Skalar IRIS System as positioned on a subject (http://www2.le.ac.uk/departments/psychology/research/facilities)*

Wilkinson Infrared Eye-Tracker

Waters and Underwood (1998) used the same apparatus as in a previous study (Wilkinson, 1979). The eye-tracker uses infrared light, reflected from the cornea of the eye “onto a photoelectric matrix” (p. 49) to distinguish the horizontal eye location. According to Wilkinson (1979), the instrument does not record the viewed stimulus but the “reflex position” (p. 429), though its precision is approximately one character space; with the eye location sampled every 4 milliseconds (Waters & Underwood, 1998). The authors describe how the instrument is situated on the subject: “the eye-movement recorder, mounted on a frame of eye-glasses, is attached and fastened to the subject's head. Head movements are restricted, but not eliminated, by use of a chin-rest and head-frame” (p. 49)

Electrooculography

Electrooculography (EOG) is the "electrophysiology of the outer retina in dark and light" (Marmor et al., 2011, p. 2). Specifically, it is the measure of variations of electrical potential through a part of the retina called the retinal pigment epithelium, while this part of the eye adjusts to intervals of darkness and light. In other words, the EOG instrument calculates the voltage between the anterior and posterior of the eye with moderate spatial and temporal resolution (http//www.dizziness-and-balance.com). The discovery of a voltage difference in the eye occurred in the mid-19th-century. Further development in the field continued in 1878, when the importance of the retinal pigment epithelium for voltage was established. In the 1940s, the changes in the eye caused by light were first recorded (Arden, 2006). According to Coupland (2006), the clinical use of EOG was initially "to indirectly measure the standing potential of the eye" (p. 252). By the mid-20th-century, researchers utilized EOG voltage to measure eye movement (Arden, 2006).

For EOG, electrodes are applied to the skin near the eye, at the outer and inner canthi of the eye, with another electrode placed elsewhere, such as the forehead, and calibration is required (http://www.dizziness-and-balance.com; Marmor et al., 2011). According to Marmor and colleagues, subjects are required to either have dilated pupils or the laboratory setting should enable consistent illumination for the retinas. EOG has been used for music reading research, as described in this chapter (Kopiez and Galley, 2002). Recently, electrooculography has been used with an electroencephalogram (EEG) in order to measure eye movement in combination with electrical activity of the brain. According to Jiang and colleagues (2014), "the integration of different electrophysiological signals into an electroencephalogram (EEG) has become an effective approach" (p. 2919). In the chapter concerning neurological assessment of music reading, the following studies employ EOG with EEG methods, and are described accordingly: Schön and Besson (2002; 2005), Gunter, Schmidt, and Besson (2003), and Wong (2010).

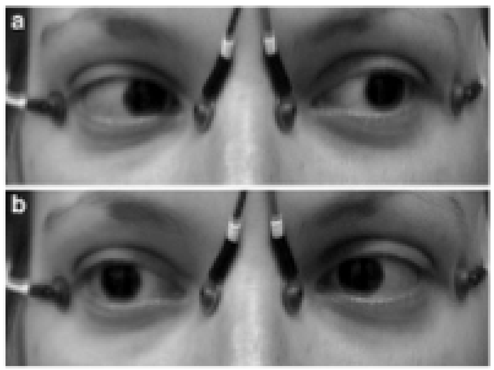


Figure 10. *Electrooculography electrodes positioned around the eyes (Marmor et al., 2011, p. 2).*

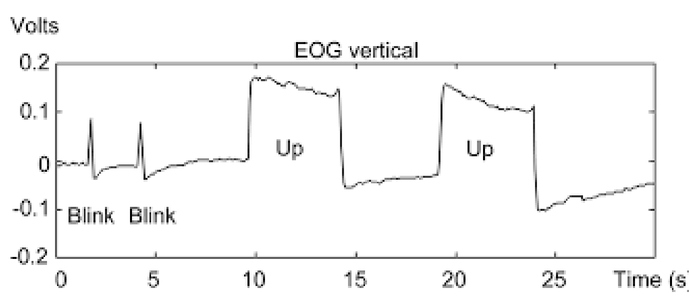


Figure 11. *An example of a Vertical EOG Scan (Barea et al., 2011, p. 318).*

4.2. Music Reading Studies Using Eye Movement Methods

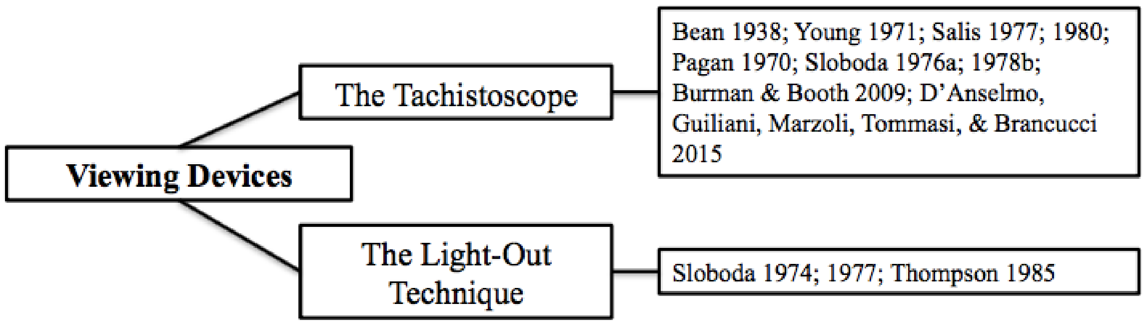


Figure 12. *Viewing devices with subcategories, and studies.*

4.2.1. Viewing Devices

The Tachistoscope

Bean, 1938	
Type	Cognitive Process
Apparatus	Tachistoscope
Production Method	Piano
Participants	50 (exp. 1), with 3 children aged 10-11, and a wide range of abilities; 24 (exp. 2), with 2 children aged 13-14, and all technically advanced
Stimuli	Author-composed melodic excerpts
Pre- and Post-Tests	No (exp. 1); Yes (exp. 2)
Control Group	No
Measurement	Eye Movement

An early and foundational study of music reading by Bean (1938) utilized the tachistoscope, where it aided in measuring note and rhythm accuracy and tempo consistency during performance. The stimuli, which comprised musical excerpts of melodic, polyphonic, or harmonic notation, were viewed for an average of .185 of a second, with the objective to view the entire stimulus with one fixation. The subject sat at the piano with the tachistoscope secured on top of the instrument and the viewing area directly in front of the subject. After viewing the stimulus, the subject performed the notation. Bean writes, "reading in the Twin Tachistoscope, the performer has no opportunity to anticipate the direction of motion of the music before he actually fixates it" (p. 15). This experiment found that effective or skilled reading of musical notation was the recognition of at least three-note patterns.

Young, 1971	
Type	Cognitive Process
Apparatus	Tachistoscope
Production Method	Piano
Participants	17, successful piano readers
Stimuli	Pre-Test: Scale and arpeggio, Sight-reading: 2 musical excerpts, Eye-movement tests: 8 Author-composed
Pre- and Post-Tests	Yes
Control Group	No
Measurement	Eye Movement

Young (1971) measured eye-movement patterns of pianists using a tachistoscope for presentation and an eye-movement camera for recording. Data from the chordal stimuli were analyzed and Young found that the eye moved in a 'zigzag' pattern on the chord and to either side of the notation. Strong readers fixated more often, with an average of two fixations per chord, and consistently fixated the bass clef.

Salis, 1977; 1980 (Phase III)	
Type	Cognitive Process
Apparatus	Tachistoscope
Production Method	Piano
Participants	26, university keyboard students, professional musicians, accompanists
Stimuli	45 experts of standard piano repertoire from 5 styles
Pre- and Post-Tests	No
Control Group	No
Measurement	Eye Movement

Salis (1977; 1980) used the tachistoscope for subjects to view dot patterns and chords. With the dot patterns, they were to simply count the number of dots and respond verbally. For the notation, subjects viewed and performed the chords, playing as many notes as possible until the experimenter ended that particular trial. Salis found that all musicians group notational stimuli, though better readers had efficient execution and related the groupings of material to each other.

Pagan, 1970	
Type	Cognitive Process
Apparatus	Tachistoscope
Production Method	Organ
Participants	12 (pilot), graduate students; 79 (main), undergraduate students
Stimuli	Pilot study and Form II of test: Author-composed chords
Pre- and Post-Tests	No
Control Group	No
Measurement	Aural-Visual

Pagan (1970) used the tachistoscope in order to develop a test to measure the formation of aural images. The purpose of the device was for controlling the presentation of the stimulus, thus four-note chords were projected with the tachistoscope at a set speed and duration, followed by an aural stimulus. Subjects had to process the notation to compare it with the aural stimulus. Pagan found that subjects with keyboard training were most successful with the task, but also that the test had the appropriate validity and reliability.

Sloboda, 1976a	
Type	Cognitive Process
Apparatus	Tachistoscope
Production Method	None
Participants	8 (exp. 1, 2); 10 (exp. 3, 4), all either musically experienced or naïve university students
Stimuli	6 cards with musical stave of 5 and 1-6 note heads (no rhythmic value)
Pre- and Post-Tests	No
Control Group	No
Measurement	Visual Coding

Sloboda, who helped to establish music reading research, adapted many text-reading paradigms, and used the tachistoscope in two of his studies. His 1976a study was comprised of four experiments. In the first, Sloboda had subjects view notation for durations of 20 milliseconds or 2 seconds and then measured their ability for recall. The stimuli were presented with a two-field tachistoscope. Sloboda found that musicians can recall up to six notes, and that they use two coding systems: a visual code is employed with the brief duration of 20 milliseconds, and a nonvisual code is employed with longer durations of 2 seconds. In the second experiment, a three-field tachistoscope was used, which held a "fixation stave", a card with three notes, and a "visual mask" containing dots and lines to muddle the stimulus. The stimulus was presented at 6 different durations before the mask also appeared, and subjects were to record the three notes at that point. It was found that musicians must have at least 100 milliseconds to code the stimulus. The final experiment used a tachistoscope to measure how an interference stimulus affects memory and how it is coded, compared with notational stimuli. The interference stimuli were paired letters or tone sequences, respectively, with three conditions using visual and auditory stimuli. Sloboda did not find reciprocal interference, and remarked that the coding of visual notational is either not by name or pitch, or that musicians are able to use the same code with another activity simultaneously. The best results were yielded when subjects viewed the notation for at least 100 milliseconds.

Sloboda, 1978b	
Type	Cognitive Process
Apparatus	Tachistoscope
Production Method	None
Participants	12, musically accomplished or musically illiterate university students
Stimuli	144 stimulus cards with notes on a staff. 48 of the cards were according to each of the 3 experimental conditions.
Pre- and Post-Tests	No
Control Group	No
Measurement	Visual Coding

Sloboda employed a 2-field tachistoscope in his 1978b study, with an adaptation of the task from the 1976a experiments. The objective was to better understand how the six musicians and six nonmusicians processed and retained notational information; specifically with patterns of one, two, or three notes in relation to the staff and to each other, respectively. Sloboda measured how a musician or nonmusician estimates the position of a note, the line of two notes, or "absolute contour" (p. 325), and the form of three notes, or "relative contour" (p. 325). Stimulus cards containing notation on two staves were assigned to one of three conditions. Conditions A and B had four notes arbitrarily placed on each staff. Condition C had four notes in a specific configuration. Subjects viewed the stimuli for 50 milliseconds and responded by transcribing the pattern to the best of their ability on an answer sheet. The results from this experiment confirmed the 1976a results: "musicians are not superior to novices at reporting briefly exposed pitch notation when an exact-accuracy measure of performance is taken" (p. 330). However, musicians could recall more information of the relative position of neighbouring notes. Kornicke (1992) summarizes that "this supports the theory that, given exposures of given duration, global (overall contour) information precedes detailed information (i.e., specific notes)" (p. 36).

Burman & Booth, 2009	
Type	Cognitive Process
Apparatus	Tachistoscope
Production Method	None
Participants	21, all adults either proficient or less-skilled
Stimuli	Author-composed for "Easy" passage, with 29 measures of single melodic line in D Major, "Difficult" passage was unfamiliar (Liszt's Etude No. 5) with 15 measures, Experimental Stimuli: Single melodic line from rehearsed passage was target sequence, consisting of 18 notes of equal duration, For each stimuli length, 4 variant sequences were used.
Pre- and Post-Tests	No
Control Group	No
Measurement	Perceptual Span

Burman and Booth (2009) measured the perceptual span with regard to the effects of rehearsal and ability. Their study included an error-detection task where stimuli were presented tachistoscopically. Subjects had to determine whether the presented musical excerpt matched the original notation previously rehearsed. Burman and Booth found that perceptual span is affected by how familiar expert musicians are with the notational patterns, but that the perceptual span was consistent to be between 4 and 5 notes, with slight increases in the span when music was rehearsed.

D'Anselmo, Guiliani, Marzoli, Tommasi, & Brancucci, 2015	
Type	Cognitive Process
Apparatus	Tachistoscope
Production Method	Piano
Participants	46, skilled pianists
Stimuli	1. Single note or chord in upper and lower staff (2-hands); 2. Note or chord only in lower staff (LH); 3. Note or chord only in upper staff (RH).
Pre- and Post-Tests	No
Control Group	No
Measurement	Perception

A recent study by D'Anselmo, Guiliani, Marzoli, Tommasi, and Brancucci (2015) also used the tachistoscope to study piano sight reading effects on perception and early processing, specifically with the two hemispheres of the brain. They measured perceptual contribution and motor response separately by comparing a variety of conditions: musical notation, verbal notation, left hand, right hand, and both hands. Chords were chosen for the musical stimuli, presented in one or both clefs, while verbal stimuli included the note name and modality written in capital letters. An advantage of the tachistoscope is the ability to present in two visual fields, and as such stimuli were presented in either the left or right visual field. After viewing the stimulus, the subject was to perform it immediately. The authors found that subjects responded faster with stimuli performed ipsilaterally, and that subjects showed a preference for the right visual field with single-handed stimuli. It was interesting to note that notation in the bass clef limited performance, especially when stimuli was shown in the left visual field. D'Anselmo and colleagues concluded that the hemispheric pattern configurations are dependent on the way the stimulus is coded as well as the type of motor response.

The Light-Out Technique

Sloboda, 1974	
Type	Cognitive Process
Apparatus	The Light-Out Technique
Production Method	Piano
Participants	10, music students of varying age and ability
Stimuli	15 unfamiliar English and French popular and folk melodies of the 1800s.
Pre- and Post-Tests	No
Control Group	No
Measurement	Eye-Hand Span

Sloboda (1974) used single-line melodies projected onto a screen, which were removed at a time unknown to the subject. The subject was instructed to play as many notes as possible without the visual aid, and the eye-hand span was measured by determining the number of correctly played notes after the slide's removal. Sloboda found that subjects' spans ranged from 4 to 7 notes, and that pianists use coding strategies to perceive visual input.

Sloboda, 1977	
Type	Cognitive Process
Apparatus	The Light-Out Technique
Production Method	Organ
Participants	6, all proficient
Stimuli	8-phrase sequences: a. Simple diatonic melody in folk or hymn melody style, with each phrase boundary marked by presence of a cadence and a relatively long interphrase space (physical and structural), b. Sequence identical to first in rhythm (same spacing), but with pitches assigned to values which made harmonic nonsense of sequence (physical), c. Sequence with comparable harmonic structure to a. but without interphrase spaces (structural), d. Sequence identical to c. in rhythm but with pitches assigned to value which made harmonic nonsense of sequence (no phrase markers)
Pre- and Post-Tests	No
Control Group	No
Measurement	Eye-Hand Span

Sloboda based his 1977 study on a text-reading paradigm to determine the effect of physical and structural markers of musical notation on the eye-hand span. With the same method of removing visual stimuli at an unknown time, he found that subjects played to a phrase boundary with either physical or structural markers, but also that structural markers increased the span.

Thompson, 1985	
Type	Cognitive Process
Apparatus	The Light-Out Technique
Production Method	Instrumental
Participants	30, aged 17-31
Stimuli	WFPS (Watkins & Farnum, 1954) Form A, Sight-reading: Randomly-generated 24-measure piece, Eye-performance Span: 7 selections from Gekeler (1969); Music Encoding Test: Slide of 2, 4, 6, 8 notes in treble clef, from Gekeler (1969) and Voxman (1954), Letter Encoding Test: Text
Pre- and Post-Tests	No
Control Group	No
Measurement	Eye-Hand Span

Thompson (1985) also measured the eye-hand span, labelling it the 'eye-performance span'. Based on Sloboda's 1974 study, music was removed at a point unknown to the subject. The subjects were to continue to play, with certainty, as many notes as possible beyond the slide's removal. Thompson was able to correlate superior sight reading skill with increased eye performance span and efficient encoding of notation. He proposed a three-stage music reading process, which occurs simultaneously: notation is encoded, which is converted into a directive for a motor response, and then executed.

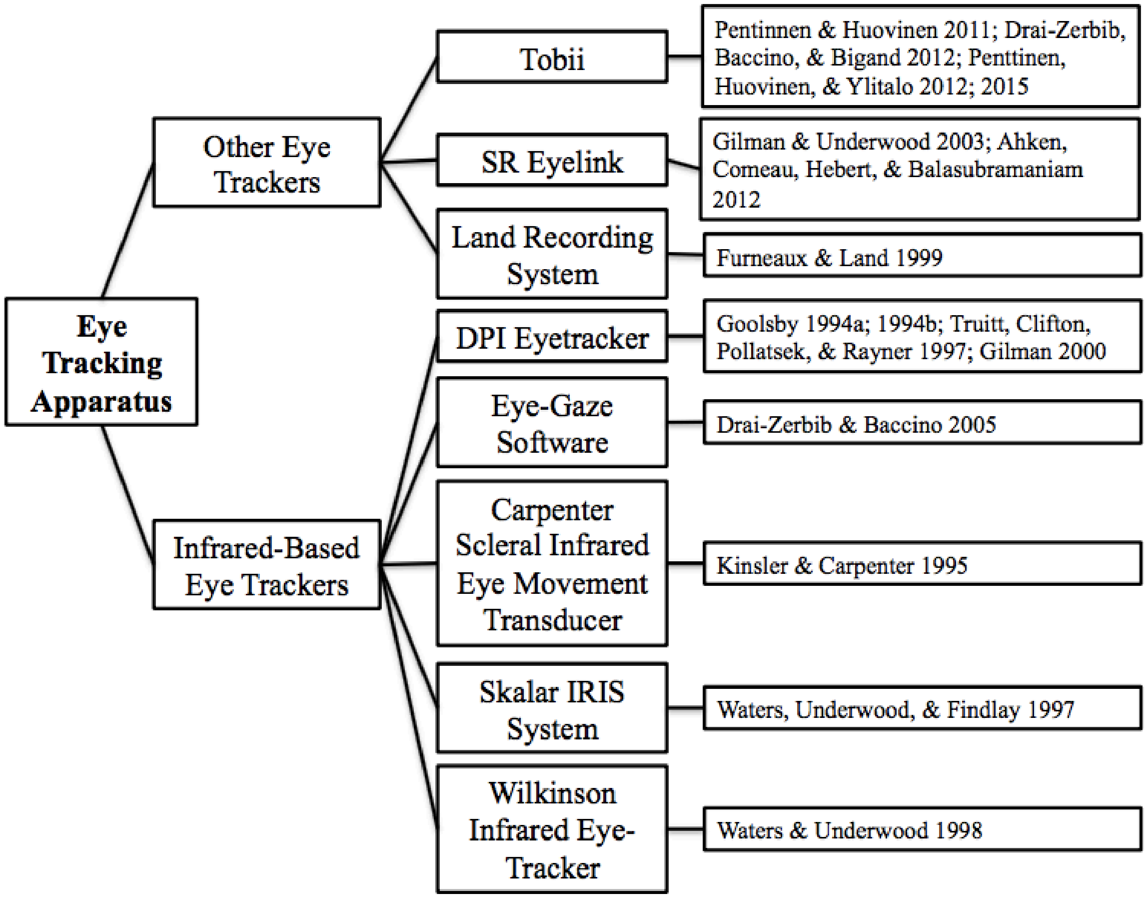


Figure 13. *Eye-tracking apparatus, subcategories, and studies.*

4.2.2. Eye Tracking Devices

Various Eye Trackers

The Tobii 1750 Eye Tracker/Tobii TX300 Eye Tracker

Pentinnen & Huovinen, 2011	
Type	Instructional
Apparatus	The Tobii Eye-Tracker
Production Method	Instrumental (with piano)
Participants	30, university education students
Stimuli	12 5-measure author-composed melodies in C major.
Pre- and Post-Tests	Yes
Control Group	Yes
Measurement	Eye Movement
Duration	4 8-week periods

Penttinen and Huovinen (2011) examined music reading development with adult subjects. As the subjects participated in music training through group piano lessons, their fixation times were routinely measured with the eye tracking apparatus. The authors found that the novice group of subjects displayed improvement, particularly with increased metrical sensitivity and by conforming to intervals, shown by more consistent fixation durations.

Pentinnen, Huovinen, & Ylitalo, 2012	
Type	Cognitive Process
Apparatus	The Tobii Eye-Tracker
Production Method	Piano
Participants	5 (pilot), with extensive musical training; 34 (main), education students minoring in music or musicians from an arts academy
Stimuli	"Mary Had a Little Lamb", in key of C Major, with variations of the notated melody (exp. 1, 2)
Pre- and Post-Tests	Longitudinal (3 data collecitons for pilot); No (main)
Control Group	No
Measurement	Eye Movement

Two experiments by Penttinen, Huovinen, and Ylitalo (2012) examined eye movement while subjects performed a familiar folk tune on the piano, followed by variations of the tune. The objective of the study was to better understand how the eyes respond to unexpected melodic events. For the first experiment, the subjects were skilled at music reading and the authors found that though the altered measures of the variations did not demand more attention, 50% of incoming saccades occurred between the preceding measures to the altered measures. Therefore, unexpected melodic events altered visual processing. The second experiment, examining a larger number of musically-literate subjects, confirmed the prior experiment's result, with additional findings of decreased fixation time and a decrease of incoming saccades for the measure following the altered measure.

Drai-Zerbib, Baccino, & Bigand, 2012	
Type	Cognitive Process
Apparatus	The Tobii Eye-Tracker
Production Method	Piano
Participants	17, students from National Conservatory of Nice, France
Stimuli	36 piano excerpts from classical tonal repertoire, with each excerpt having 3 versions for each fingering condition: Difficult, easy, or no fingering
Pre- and Post-Tests	No
Control Group	No
Measurement	Eye Movement

Drai-Zerbib, Baccino, and Bigand (2012) employed a Tobii Eye Tracker to measure eye movement for nine specific areas of a score: four measures of notation in each clef, as well as the clef area. Of these measures, the authors collected data on fixation duration, first- and second-pass fixations, and the probability of re-fixation. They also measured errors. The subjects, all proficient or expert pianists, selectively heard the stimulus played, then read the notation, and finally performed the stimulus. Stimuli consisted of piano excerpts notated with easy fingering, difficult fingering, or no fingering. Drai-Zerbib and colleagues hypothesized that pianists are able to process musical information by visual or auditory means, and found that an auditory excerpt allowed for the right hand, or treble clef, to be processed before its visual presentation. Thus, the left hand notation is processed while the stimulus is viewed for the first time, but without a previous auditory stimulus, the right hand notation received more attention.

Pentinnen, Huovinen, & Ylitalo, 2015	
Type	Cognitive Process
Apparatus	The Tobii Eye-Tracker
Production Method	Piano
Participants	38, education university students and music students of an arts academy
Stimuli	The original version of "Mary Had a Little Lamb", notated in C major, as well as 2 slightly altered versions, where target bar did not greatly deviate in visual appearance from original
Pre- and Post-Tests	No
Control Group	No
Measurement	Eye-Hand Span

Penttinen, Huovinen, and Ylitalo (2015) measured eye-hand span and gaze activity with a Tobii eye tracker. They studied how a musical score is processed, particularly with temporal regulation. As such, ‘gaze activity’ is defined as, “the number of inspected visual beats between two beat onsets” (p. 39). Penttinen and colleagues also wanted to explore how these eye movements are affected by changes in the melody line or by musical education and experience. To further explain their hypothesis regarding gaze activity:

> More experienced musicians, who assumingly operate with shorter fixation durations and thus with more frequent fixations, generally demonstrate more gaze activity than less experienced musicians. We also expect that gaze activity is affected by rhythmic patterns and the presence of melodic alterations in the following ways: firstly, melodic alterations of familiar musical material should require more visual attention and thus locally limit gaze activity, and secondly, longer notes (here, quarter notes) should leave more time for broader advance inspections, especially in the case of more experienced musicians (p. 39).

The original score of a familiar folk tune and modified versions of the tune, similar in appearance to the original notation, were divided into twelve ‘areas of interest’ for measurement. Each of these areas was matched with a quarter note beat in each stimulus. The subjects, all adult pianists, were required to perform the original tune in the right hand, in time with a metronome. After a brief practice phase where eye movements were recorded, the final test phase required subjects to play the modified stimuli at the previous tempo. Penttinen and colleagues found that the eye-hand span length of 1 second is consistent with existing literature, and that an expertise effect of shorter fixations occurred even under temporal control. With eighth-note patterns, gaze activity increased for all subjects, but there was an expertise effect of a regular increased eye-hand span and gaze activity. Finally, the altered measures tended to cause shorter eye-hand span and reduced gaze activity. This publication may challenge current beliefs about eye-movement in music reading. The authors posit that the eye-hand span is a consistent length of time, regardless of stimuli or skill-level, which "demonstrates that the eye–hand span neither directly reflects the experienced difficulty of the notated material, nor does it reflect sight-reading competence alone" (p. 48). The authors offer that the eye-hand span may be restricted by the temporal confines of short-term memory, or that it may aid in "performing the music in appropriate tempo rather than the purpose of performing at maximal capacity" (p. 48).

SR EyeLink

Gilman & Underwood, 2003	
Type	Cognitive Process
Apparatus	SR Eyelink
Production Method	Piano
Participants	40, all with Grade 8 piano (UK) or equivalent
Stimuli	32 musical phrases from Bach chorales (Riemenschneider, 1941 (Ed.)). Tenor parts were removed, For error detection task: phrases had 1-4 notational "errors".
Pre- and Post-Tests	Yes
Control Group	No
Measurement	Perceptual Span

Gilman and Underwood (2003) used the moving-window technique to measure the eye movements and perceptual spans of pianists, among other measures. All subjects participated in sight-reading, transposition, and error-detection tasks under four conditions: 1-, 2-, and 4-beat 'gaze-contingent' windows, and no window. Their key findings were consistent with a previous study by Truitt and colleagues (1997): all music readers, regardless of skill-level, have similar perceptual spans, and perceptual span is not affected by increasing visual information. It is interesting to mention that the two studies employing the moving-window paradigm achieved similar results, though the stimuli used by Truitt and colleagues were melodically- and rhythmically-simple excerpts using a 5-note range and played with one hand, while the stimuli of Gilman and Underwood's study were three-voice chordal excerpts played with both hands together. Gilman and Underwood (2003) offer two explanations: that there is no increase in perceptual span with the complex chordal harmony because the cognitive load is not enlarged and thus the amount of foveal material remains, or that "musicians could have made use of a larger perceptual span when reading a single line melody, but didn't because doing so would have placed a high demand on working memory" (p. 226). The farther parafoveal vision is encompassed in the perceptual span, the more working memory is employed to remember notation in successive fixations. However, Gilman and Underwood conclude that the findings of both studies suggest that one's perceptual span remains constant.

Ahken, Comeau, Hebert, & Balasubramaniam, 2012	
Type	Cognitive Process
Apparatus	SR Eyelink
Production Method	Piano
Participants	18, experienced pianists
Stimuli	Linguistic: 20 short sentences, half of which included syntactic anomaly (either progressive "-ing" affix or past-tense inflection) (Taken from Braze et al., 2002) Musical: Graduate-student composed, 16 short musical phrases in groups of 4. Each phrase had 5, 6, or 7 bars. Half of the sequences were syntactically congruent. Written in 2 forms: key signature OR with accidentals.
Pre- and Post-Tests	No
Control Group	No
Measurement	Eye Movement

Ahken, Comeau, Hébert, and Balasubramaniam (2012) employed the SR Eyelink II to explore eye movement during visual processing with syntactic incongruities. Using both music notation and written text, the authors investigated the effects of syntactic incongruity, and with music: the effect of key signature for syntactic comprehension. Ahken and colleagues measured total reading time as well as eye fixations, fixation duration, and regressions. The musical stimuli were divided into 'regions of interest', to study eye movements for each measure of music. Of the stimuli, half were syntactically congruent, and the score had either a key signature or accidentals. Subjects alternated performing the musical stimuli with reading text, and the authors found that with incongruent stimuli, there were more fixations, and both fixation and trial durations were increased. According to Ahken and colleagues, "Our study therefore provides the first evidence that eye-movement patterns are sensitive to music-syntactic incongruities" (p. 23). They also found that music notation is processed at the same rate, whether there is a key signature or the use of accidentals.

Land Recording System

Furneaux & Land, 1999	
Type	Cognitive Process
Apparatus	The Land Recording System
Production Method	Piano
Participants	8 adult pianists: 3 'novices' (ABRSM Grade 3-4), 3 'intermediates' (ABRSM Grade 6-7), 2 professionals (accompanists)
Stimuli	5 excerpts of musical scores published under a particular grade standard, with each assigned 2 tempos: "very slow" and "very fast".
Pre- and Post-Tests	No
Control Group	No
Measurement	Eye-Hand Span

Furneaux and Land (1999) used an eye tracking device that measures eye fixations and saccades to determine the eye-hand span. They described measurement of the eye-hand span by two means: the number of notes between fixation and execution, or the elapsed time between fixation and motor response. They consider the former the 'note index' and the latter the 'time index'. Furneaux and Land describe the process in the interim:

> Many things happen during this separation. Printed material must be recognized, deciphered and processed. This information must then be stored within an internal buffer, and all material to be performed simultaneously must be similarly processed. The concurrent material must then be reassembled, probably within the buffer, before motor output can occur. In addition, the information must be stored until performance has reached the appropriate part of the sequence. Therefore, the length of time from eye fixation to hand performance is a measure of the total time involved in processing and storage, prior to output (p. 2435).

Before this point, the time interval of the eye-hand span had not been measured. Furneaux and Land obtained three key findings: improvement in ability correlates with an increase in note index of the eye-hand span. The mean time index, which is an average of one second, is not affected by skill level. Finally, when the tempo of performance was increased, the time index decreased, which was consistent for all subjects. The authors suggest that the increase in the note index of the eye-hand span is due to an efficient storage mechanism, allow for a greater amount of notation to be viewed and stored before performance. They write, "This does not necessarily mean that better readers...have simply developed larger storage buffers. An alternative explanation is that rather than reading and processing individual notes, professionals are able to chunk several notes together and process them as a single unit of information, enabling them to store more in a buffer of similar capacity" (p. 2439).

Infrared-Based Eye Trackers

The Dual-Purkinje-Image Eyetracker

Goolsby, 1994a	
Type	Cognitive Process
Apparatus	The DPI Eyetracker
Production Method	Voice
Participants	24, graduate music students
Stimuli	Four melodies had various levels of notational complexity, from *Solfege des Solfeges* (Danhauser, Lemoine, & Livigna, 1910)
Pre- and Post-Tests	No
Control Group	No
Measurement	Eye Movement

Goolsby was the first to employ eye-tracking technology that continues to be used in current research: The Dual-Purkinje-Image (DPI) Eyetracker. In his 1994a study, he examined six types of eye movements: the number of progressive fixations, progressive fixation duration, progressive saccade length, the number of regressive fixations, regressive fixation duration, and regressive saccade length. The twenty-four subjects were selected by the results of an initial sight reading test (the subjects with the twelve highest scores and the twelve lowest scores participated in the experiment). The subjects were required to sight-sing four melodies of varying complexity. They sang each melody three times: once *prima vista*, a second time immediately following, and a third time after the subjects were given time for study and practice. Goolsby found that notational complexity reduced eye movement and that skilled readers look ahead in the music, and then back to where they are performing, resulting in fewer but longer fixations. He also measured eye movement in relation to perceptual processing, and found differences in eye-movement among varying skill levels, and that notational spacing affected all subjects.

Goolsby, 1994b	
Type	Cognitive Process
Apparatus	The DPI Eyetracker
Production Method	Voice
Participants	2, representing 2 distinct levels of ability
Stimuli	3 melodies varying in notational complexity, but with the same physical dimensions.
Pre- and Post-Tests	No
Control Group	No
Measurement	Eye Movement

In a subsequent study, Goolsby (1994b) further examined eye movement differences among subjects to better understand the cognitive foundation for eye motion control. The DPI Eyetracker was used to measure fixations and saccades, and Goolsby found expertise effects for progressive saccades, multiple regressive saccades, and the suggestion of notation processing in parafoveal vision. According to Madell and Hébert (2008), Goolsby's studies are foundational for eye movement research in the field of music reading.

Truitt, Clifton, Pollatsek, & Rayner, 1997	
Type	Cognitive Process
Apparatus	The DPI Eyetracker
Production Method	Piano
Participants	8, average to above-average ability
Stimuli	32 simple melodies, played with one hand, from beginning of *Mikrokosmos, Vol. 1* (Bartók, 1940). Each melody divided into non-overlapping 2-beat regions; 2 regions per measure. No-window condition: Entire melody displayed on screen throughout the trial. 2-beat, 4-beat, 6-beat conditions: Fragments of the melody were added to the display as eyes moved through the region.
Pre- and Post-Tests	No
Control Group	No
Measurement	Perceptual Span

Truitt and colleagues (1997) employed four conditions with the moving window technique to measure the perceptual span: 2-, 4-, and 6-beat melodic portions of the stimulus displayed as the eyes followed the stimulus, and no window. The data reveals that most subjects, regardless of skill level, fixated and processed a single note in advance of a motor response and at most 3-4 beats of music in advance of the performed note. In general, the perceptual span is slight; between two and four beats in length, with expert readers possibly viewing a span of five notes ahead of the hand. However, Truitt and colleagues propose that musicians may be able to obtain useful information from a larger area if the means for the task does not require a motor response, and give the example of visually detecting notational errors on a score without accompanied performance. They suggest that musicians "are seduced by the illusion that detail can be extracted from a wider region of the parafovea and periphery than is actually possible" (p. 160).

Gilman, 2000 (Experiments 9 and 10)	
Type	Cognitive Process
Apparatus	The DPI Eyetracker
Production Method	Piano
Participants	20 (exp. 9), all passed Grade 8 or equivalent, and half considered "good"; 16 (exp. 10), 8 "good" and 8 "poor"
Stimuli	Triads (exp. 9); Composed phrases from Bach's chorales (Riemenschneider, 1941), and edited by experimenter so that all comprised 35 notes in total. For each of the 224 phrases, a "nonidentical twin" was created.
Pre- and Post-Tests	No
Control Group	No
Measurement	Eye Movement (exp. 9); Parafoveal Vision (exp. 9 and 10)

Objectives of Gilman's (2000) thesis were to investigate how notation viewed in the parafoveal region is processed, and to what extent the type or types of material is processed during preview. Gilman compares her data with Kinsler and Carpenter's (1995) proposed reading model, with a particular focus on eccentricity. Eccentricity is defined as the area from the point of fixation to the object (Gilchrist, 2011) and is shown in degrees. Using the apparatus for two of the ten experiments, Gilman measured the decrease of comprehension with increased eccentricity, using root position and inverted triads of early-acquired keys. Gilman found that strong readers have the ability to encode notation from extrafoveal vision, and that some elements of notation, such as the form of the music, is not as affected by eccentricity than individual notes are. Gilman writes, “this finding is important as it implies that any model of sight-reading should take into account that notes presented near to the fovea are processed to a different level than notes presented further from the fovea” (p. 270). In a subsequent experiment, Gilman explored the boundaries of extrafoveal-processed material, replicating the previous experiment with the use of eye-tracking apparatus. The results were consistent with the previous data, but also indicated that greater eccentricity results in diminished performance. In the final experiment of the Gilman (2000) study, the effect of expectancy on parafoveal preview was studied. The stimuli consisted of phrases adapted from Bach chorales where the notation was either proximal (notes closer together) or distal (notes farther apart). Proximal stimuli allow for greater preview of upcoming notation in the parafoveal vision, and Gilman found that this resulted in fewer, yet longer fixations, with better performance results. Gilman writes, “it is evident that we use extrafoveal information to facilitate processing” (p. 241).

Eye-Gaze Software

Drai-Zerbib & Baccino, 2005	
Type	Cognitive Process
Apparatus	Eye-Gaze Software
Production Method	Piano
Participants	27, expert and non-expert pianists
Stimuli	16 musical partitions, containing scores from Czerny, Bartók, Scarlatti, or author-composed scores.
Pre- and Post-Tests	No
Control Group	No
Measurement	Eye Movement

Drai-Zerbib and Baccino's 2005 study examined recall of musical excerpts processed by visual or auditory means, with select stimuli containing slurs for articulation. Using *Eye-Gaze* software on a computer monitor, the researchers measured eye movement of the stimuli from the dominant eye, including fixation, saccades, and the temporal characteristics of these movements. Drai-Zerbib and Baccino found expertise effects regarding the region of the score fixated and significant interactions, including a twofold interaction between viewed areas, the presence of slurs, and expertise.

Carpenter Scleral Infrared Eye Movement Transducer

Kinsler & Carpenter, 1995	
Type	Cognitive Process
Apparatus	Carpenter Scleral Infrared Eye Movement Transducer
Production Method	None (Tapping Response)
Participants	4, all competent musicians. 2 were informed of the experiment and 2 were not.
Stimuli	Single line of simplified music with rhythmic information only.
Pre- and Post-Tests	No
Control Group	No
Measurement	Parafoveal Vision

Kinsler & Carpenter (1995) measured eye movements while subjects remained as motionless as possible. This was achieved by requiring the subjects to respond to the stimuli by simply tapping the viewed rhythm onto a microphone. Kinsler and Carpenter found that expert musicians would fixate approximately midway through the last measure of a score, but would play to the end of the score. They concluded that parafoveal vision allows for continued performance, as an advanced visual system is still processing the information seen extrafoveally. From the collected data, the authors proposed a reading model employing both oculomotor and perceptual operations, where one scans and processes the stimulus as far as there is recognized information either by clear visibility or predictability, or due to necessity, such as in the case of sight reading. This scan includes information received from parafoveal vision.

Skalar IRIS System

Waters, Underwood, & Findlay, 1997	
Type	Cognitive Process
Apparatus	Skalar IRIS System
Production Method	None
Participants	Unknown (pilot); 24 (main), 3 groups: 8 university music students, 8 psychology students with music experience on monophonic instrument, 8 nonmusicians (ages 18-29 years)
Stimuli	60 generated melodies
Pre- and Post-Tests	No
Control Group	No
Measurement	Perception

Waters, Underwood, and Findlay (1997) explored the “perceptual processing of music notation” (p. 486) with the use of a pattern-matching model in pilot and main studies. Their objective was to confirm that expert readers have the ability to process groups of notes quickly. Subjects were to decide if the two musical sequences viewed were the ‘same’ or ‘different’. The stimuli consisted of either temporally coherent, pitch coherent; temporally coherent, pitch randomized; temporally randomized, pitch coherent; or temporally randomized, pitch randomized sequences of ten notes. The pilot study did not measure eye movement but validated the pattern-matching task, with findings of expertise effects for reaction times, particularly with the ‘different’ trials. For the main study, response times and choice were measured, as well as the horizontal and vertical signal from the eyes, respectively, by differential limbus reflection with the Skalar IRIS System. For this study, there were three groups of subjects: music students (experts), psychology students with musical background (novices), and nonmusicians. Similar stimuli was composed, but with the additional factor of having the notes close or more distant from each other. The expert and novice subjects had faster reaction times and heightened sensitivity to temporal organization. All subjects made more choice errors when the notation was close together, but processed the notation faster. For eye movement, Waters and Colleagues found that experts process longer units of notation, and that both expert and novice subjects viewed the stimuli for shorter periods, with fewer and shorter fixations.

Wilkinson Infrared Eye-Tracker

Waters & Underwood, 1998	
Type	Cognitive Process
Apparatus	Wilkinson Infrared Eye-Tracker
Production Method	None
Participants	22
Stimuli	Author-composed excerpts: “tonally simple, visually simple”, “tonally simple, visually complex”, “tonally complex, visually simple”, and “tonally complex, visually complex”.
Pre- and Post-Tests	No
Control Group	No
Measurement	Eye Movement

The subjects in Waters and Underwood’s (1998) study completed a same-different task with short excerpts of notation, while the eye-tracking device measured eye movements. The authors found expertise effects with accuracy and speed, especially with tonally simple stimuli. The strong readers also employed more fixations with the first stimulus of the pair, but fewer fixations before giving a response. This data contradicts the results from Goolsby (1994a), but it may be due to differing stimuli and methods of the studies.

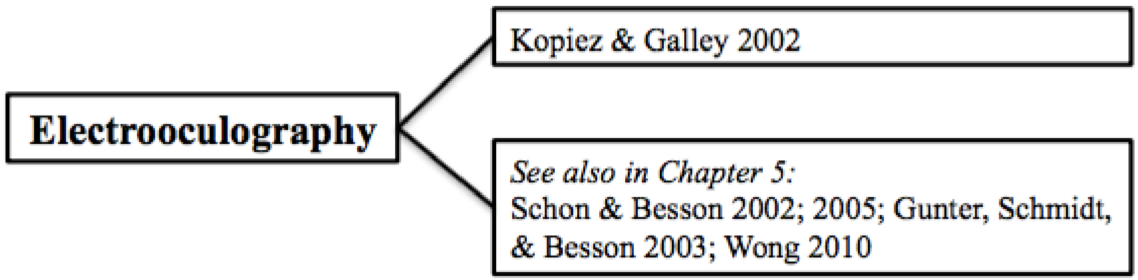


Figure 14. *Electrooculography, with studies.*

Electrooculography

Kopiez & Galley, 2002	
Type	Cognitive Process
Apparatus	Electrooculography
Production Method	None
Participants	262, 8 professional musicians and 254 psychology university students
Stimuli	A horizontal shifting dot ("jumping point")
Pre- and Post-Tests	No
Control Group	Yes
Measurement	Eye Movement

Kopiez and Galley (2002) utilized electrooculography (EOG) with an EOG amplifier device. For this study, subjects rested their head in place to limit head movement. The EOG amplifier measured the electrical field as it fluctuated with dipole movement of the eye; specifically, with sinusoidal, rectangular, and counter movements. For this measure, subjects were to visually track a shifting dot presented on a monitor. The authors also used EOG with closed eyes. Kopiez and Galley found distinct eye-movement for musicians, and defined it as the 'musicians' glance'. As well, expert musicians were more efficient with the processing of stimuli.

4.3. Summary

For **reading models**, there is a proposed 'parallel processing concept' for sight reading, which involves notational encoding that is converted into a motor command and executed. This three-stage process occurs simultaneously, resulting in a greater eye-performance span and efficient sight reading skill (Thompson, 1985). There is a visual code, which operates with brief viewing of a stimulus and which all musicians use, and a nonvisual code, which allows musicians to efficiently view and recall notation (Sloboda 1976a). Skilled musicians employ effective processing strategies with visual-motor tasks and possess the 'musicians' glance', which involves fewer errors, shorter reaction times, pre-emptive and faster saccades, and quicker anticipatory latencies (Kopiez & Galley, 2002). There is also a proposed oculomotor and perceptual model, which includes internalizing the area viewed for interpretation (Kinsler & Carpenter, 1995). A proposed dual coding model is applied for interval processing, in part demonstrated by the fact that poor readers are not as proficient with abstract processing and must rely on the visuo-spatial features of the notation. The subskills of sight reading, as related to eye movement, include pattern recognition, preview, and the ability to begin to process material through means of peripheral vision (Gilman, 2000).

For **general eye movement**, skilled readers exhibit longer progressive saccades and more saccades between the staves. Increased tempo decreases saccadic interval with simultaneous increase of saccadic amplitude (Kinsler & Carpenter, 1995; Goolsby, 1994a; Furneaux & Land, 1999). Longer note values or further proximity between notes results in increased fixations and longer units of view (Waters, Underwood, & Findlay, 1997; Kinsler & Carpenter, 1995). Musical context also allows for fewer and shorter fixations, and predictability results in decreased saccadic rate and amplitude (Gilman, 2000). As well, musicians display notable perceptual comprehension when notation is presented for 100 milliseconds or longer (Sloboda, 1976a).

The average eye-hand span is one second in length (Penttinen, Huovinen, & Ylitalo, 2015) and less than one beat for less-skilled readers and two beats for skilled readers (Truitt et al., 1997). However, Gilman and Underwood (2003) found a greater eye-hand span correlates with skill-level and decreases with increased cognitive load. The perceptual span is also small (Truitt, Clifton, Pollatsek, & Rayner, 1997) and unaffected by skill-level and cognitive load (Gilman & Underwood, 2003), but it may be increased by rehearsal (Burman & Booth, 2009). Gilman (2000) found that experts have a greater note identification span, which suggests that good readers have a larger perceptual span extending beyond the note identification span.

For **eye movement with score details**, the eye-hand span is affected by details in a score (Sloboda, 1974; 1977; Goolsby, 1994a). During performance, piano notation in the treble clef is consistently processed more than the bass clef, and simply listening to a score can also facilitate processing of the treble clef, often resulting in regressed eye-movement to important information, such as the key signature (Drai-Zerbib, Baccino, & Bigand, 2012). Ahken, Comeau, Hébert, and Balasubramaniam (2012) suggest that music notation is processed at the same speed whether the score is written with accidentals or a key signature, and there is an expertise effect of fewer fixations when viewing a score, no matter if articulated slurs are present or absent (Drai-Zerbib & Baccino, 2005; Waters & Underwood, 1998). However, Young (1971) found expertise effects of frequent fixations on chords, a greater number of fixations of the bass clef, with fewer fixations during preview. As well, altered notation results in more incoming saccades from a previous measure to an altered measure and shorter fixation time following an altered measure (Penttinen, Huovinen, & Ylitalo, 2012; Penttinen, Huovinen, & Ylitalo, 2015; Ahken, Comeau, Hébert, & Balasubramaniam, 2012).

There are **expertise effects** with creating aural representations (Pagan, 1970), grouping notation and relating groups with each other (Salis, 1977; 1980; Gilman, 2000), increased temporal sensitivity (Waters, Underwood, & Findlay, 1997), increased musical context and a suggestion of increased parafoveal comprehension (Gilman, 2000; Goolsby, 1994b), viewing and storing notation effectively (Furneaux & Land, 1999). Reading skill improves with training, particularly with metrical sensitivity, and a decrease in fixations of intervallic notation (Penttinen & Huovinen, 2011).

Table 14. *Synthesis of summary tables for viewing apparatus.*

	Apparatus	Production Method	Participants	Measurement
Bean 1938	Tachistoscope	Piano	50; 24 (children)	Eye Movement
Young 1971	Tachistoscope	Piano	17 (proficient)	Eye Movement
Salis 1977; 1980	Tachistoscope	Piano	26 (university; adults)	Eye Movement
Pagan 1970	Tachistoscope	Organ	12; 79 (university)	Aural-Visual
Sloboda 1976a	Tachistoscope	None	8; 8; 19; 19 (proficient/not proficient)	Visual Coding
Sloboda 1978b	Tachistoscope	Transcription	12 (proficient/not proficient)	Visual Coding
Burman & Booth 2009	Tachistoscope	None	21 (proficient/not proficient)	Perceptual Span
D'Anselmo et al. 2015	Tachistoscope	Piano	46 (proficient)	Perception
Sloboda 1974	Light-Out Technique	Piano	10 (mixed)	Eye-Hand Span
Sloboda 1977	Light-Out Technique	Organ	6 (proficient)	Eye-Hand Span
Thompson 1985	Light-Out Technique	Instrumental	30 (adults)	Eye-Hand Span

Table 15. *Synthesis of summary tables for eye-tracking apparatus.*

	Apparatus	Production Method	Participants	Measurement
Pentinnen & Huovinen 2011	Tobii Eye Tracker	Instrumental; Piano	30 (university)	Eye Movement
Pentinnen, Huovinen, & Ylitalo 2012	Tobii Eye Tracker	Piano	5; 34 (university)	Eye Movement
Drai-Zerbib et al. 2012	Tobii Eye Tracker	Piano	25 (university)	Eye Movement
Pentinnen, Huovinen, & Ylitalo 2015	Tobii Eye Tracker	Piano	38 (university)	Eye-Hand Span
Gilman & Underwood 2003	SR Eyelink Eye Tracker	Piano	40 (proficient; less proficient)	Perceptual Span
Ahken et al. 2012	SR Eyelink Eye Tracker	Piano	18 (proficient)	Eye Movements
Furneaux & Land 1999	Land Recording System Eye Tracker	Piano	8 (adults; proficient; less proficient)	Eye-Hand Span
Goolsby 1994a	DPI Eyetracker	Sight-singing	24 (university; proficient; less proficient)	Eye Movement
Goolsby 1994b	DPI Eyetracker	Sight-singing	2 (university; proficient; less proficient)	Eye Movement
Truitt et al. 1997	DPI Eyetracker	Piano	8 (proficient and very proficient)	Perceptual Span
Gilman 2000	DPI Eyetracker	Piano	20; 16 (3 levels of proficiency)	Eye Movement; Parafoveal Vision
Drai-Zerbib & Baccino 2005	Eye-Gaze Software Eye Tracker	Piano	27 (proficient; nonproficient)	Eye Movement
Kinsler & Carpenter 1995	Scleral Transducer Eye Tracker	None	4 (proficient)	Parafoveal Vision
Waters, Underwood, & Findlay 1997	Skalar IRIS System Eye Tracker	None	Unk. (proficient; nonproficient); 24 (proficient; less proficient; nonproficient)	Perception
Waters & Underwood 1998	Wilkinson Eye Tracker	None	22 (proficient; less proficient)	Eye Movement
Kopiez & Galley 2002	Electrooculography	None	262 (8 musicians; 254 nonmusicians)	EOG for Eye Movement

Table 16. *Summary of number of studies by classification.*

Measurement	**Cognitive Process**	**Instructional**	**Correlational**	**Test Validation**
Number of Studies	**26**	**1**	**0**	**0**

CHAPTER 5

5. Neurological Measurements

There are many areas of the brain activated while perceiving or performing music. An auditory or visual stimulus provides information, which is converted from sensation to electric impulses, or neural signals, to be encoded (Fitzpatrick & Mooney, 2012). The cerebral cortex, which is the outer layer of the brain, processes these stimuli. As early as infancy, and even before birth, activation to music and speech has been found in the cerebral cortex (Hodges, 2015). Among the many structures of the cortex, the ones involved in movement for music performance are: the premotor cortex that organizes and chooses movements, particularly those movements elicited by sensory cues, the motor cortex that projects to circuits in the brainstem and spinal cord, and with fine movements of the hands and fingers (White & Hall, 2012). Krings and colleagues (2000) remark that the level of involvement with complex movements is unknown, but the following areas play a role: the motor area is concerned with task sequencing, task complexity, and initiation of movement; the premotor areas display increased activation with guided movements or movements elicited by sensory stimuli; the primary sensorimotor cortex has supervisory functions, and the degree of complexity with a task activates the superior parietal areas of the cerebrum.

It is also understood that the cerebellum plays a part in motor learning and further cognitive processes, though details of its operations are still unknown (Nolte, 2010; Schlaug, 2001). However, White and Hall (2012) identify the cerebrocerebellum as having involvement with the regulation of proficient movements. Intricate spatial and temporal sequences are organized and implemented by this area of the cerebellum. As well, Nolte (2010) lists general roles of the cerebellum: aiding in the coordination of movement, including the comparison of intended limb movements to actual limb movements, orientation and posture of the body, managing smooth eye movements, modifying reflexes, and learning novel physical skills. He furthers that the vermis region of the cerebellum receives somatosensory information from parts of the visual and auditory cortex.

There is not as much known about the cerebral structures involved in music reading, though neuroscience methods allow researchers to explore how sight reading shifts from an artificial to natural process, and how specific areas of the brain are engaged for this process (Stewart et al., 2003b). According to Schlaug (2001), neurological study of musicians is advantageous to explore the cerebral properties of unique skills like sight reading. A brief discussion of neuroplasticity and critical periods for learning follows. Music reading studies will then be classified, by neuroimaging method and within the area of brain injury research. Each description contains measurement and findings, including areas of the brain involved in music reading.

For neuromusical research, studying neuroplasticity and its adaptations provides insight into how music training affects the brain and how new skills are acquired. According to Hodges (2015), 'plasticity' means that structures in the brain are impressionable. The environment and one's genetic make-up influence each other in a cooperative relationship to shape the developing brain, and specifically the circuits in the brain (Hodges, 2015; Lamantia, 2012). In fact, each brain cell, also known as a neuron, has the capability to connect with thousands of other neurons. These neuronal connections are established and retained, or, conversely, there may be no need for a connection and the connection is ended. External stimuli cause alterations or adjustments of these connections; thus, altering the brain's plasticity (Schlaug, 2001). Therefore, a particular environment with specific stimuli, such as exposure to music or musical training, results in a unique neural system for each person (Hodges, 2015). Skill acquisition, like music reading, is interesting for neuroscience researchers because most musical training begins while the brain is still able to adapt. This is why musicians can be used as a model for studying functional and structural modification of the brain (Schlaug, 2001).

With the developing brain, there are "temporal windows" (Lamantia, 2012, P. 537), when experience or skill acquisition has the greatest effect, and when stimulation is needed for natural growth (Berk, 2001). These are called critical periods. According to Lamantia (2012), "correlated patterns of activity are thought to mediate critical periods by stabilizing concurrently active synaptic connections and weakening or eliminating connections whose activity is divergent" (P. 556-557). When these critical periods end, later experience does not have the same influence on the behavior. Optimal periods, also known as sensitive periods, differ in that skills are acquired easily and rapidly, though acquisition can also happen outside of these windows at a slower pace (Hodges, 2015). Also, Lamantia (2012) states that the environment is crucial for early critical periods, while later critical periods for the acquisition of sensorimotor skill and complex behaviors are more broad and taper gradually. "The availability of instructive experiences from the environment, as well as the neural capacity to respond to them, is key for the successful completion of the critical period" (P. 539). It is generally unknown if the skill of music reading is influenced by early introduction to music and inborn ability, or acquired during critical periods when the brain is developing and maturing (Schlaug, 2001).

5.1. Neuroscience Method Descriptions

Other neurological studies use a variety of imaging techniques for specific data acquisition concerning brain structure, modification, and activation, and are appropriate for cognitive neuroscience (Carter & Shieh, 2015). Bunge and Kahn (2009) state that a significant contribution of neuroimaging is with the study of higher cognitive functions, which are the foundation for language and problem solving, among other mental functions. Carter and Shieh (2015) state that imaging brain cognition activity "has provided scientists the opportunity to correlate activity in distinct brain regions with specific mental operations, a truly remarkable achievement" (P. 1). There are two classifications for whole-brain imaging: structural imaging for the anatomy and functional imaging for the underlying physiological processes. Structural imaging can be used in conjunction with functional imaging. A brief history and explanation of the methods follow.

5.1.1. X-ray and Computerized Tomography

In the 20th-century, scientists and engineers sought to improve brain-imaging technology. The predominant method for viewing internal organs was with X-ray machinery, but this means could not distinguish or show contrast between brain substances, which are proteins and carbohydrates, fat, and salt water. As early as 1919, X-rays were enhanced with the use of dyes, when scientists would inject various fluids to provide contrast in their images (http://www.slac.stanford.edu/pubs/beamline/25/2/25-2-linton.pdf). Later, in the 1970s, computerized (axial) tomography ("CT" or "CAT" scan) was developed (http://www.slac.stanford.edu/pubs/beamline/25/2/25-2-linton.pdf; Carter & Shieh, 2015). CT was the original technique to generate an image with the aid of a computer; reforming information from many X-ray beams to produce an image (Kevles, 1997). However, X-rays emit radiation and CT scans can only produce images on a predetermined angle (Carter & Shieh, 2015).

5.1.2. Positron Emission Tomography (PET)

Positron emission tomography (PET) began with an early machine built in 1951. Scanning with positron detectors for research use flourished throughout the 1950s, though PET technology was not readily available for use until the mid-1980s. However, once this method was established, PET persisted in its importance as a research tool. PET technology shows the physiology of the brain without showing structure (Derdeyn, 2007). It involves "an unstable positron-emitting isotope...injected into a subject's carotid artery" (Carter & Shieh, 2015, P. 17). This artery provides blood for the cerebral hemisphere. With cognition or motor tasks, the activated area of the brain intakes a portion of the injected material (Comeau, 2016). PET is best utilized for studying the metabolism of neurotransmitters or other molecules as it "can provide very sensitive and quantitative measurements of the presence of any labeled molecule" (Derdeyn, 2007, P. 87). This method also determines the metabolic effect of lesions in the distal regions of the brain, making this method uniquely advantageous for these measures. However, PET is invasive and as such there cannot be multiple studies on the same subject. It may be the most expensive technology, and the resolution is not as strong as other methods (Carter &

Shieh, 2015).

Figure 15. *A PET scanner (http://bme.ucdavis.edu/cmgi/services/positron-emission-tomography-pet/)*

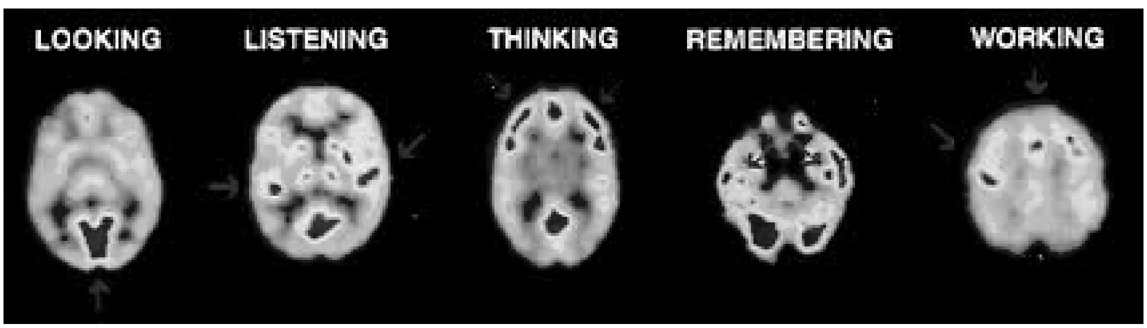


Figure 16. *An example of a PET scan (Phelps, 2016).*

5.1.3. Nuclear Magnetic Resonance (NMR), Magnetic Resonance Imaging (MRI), and Functional Magnetic Resonance Imaging (fMRI)

With the continued development of computers, and developed simultaneously with PET, nuclear magnetic resonance (NMR) technology was established, the precursor to magnetic resonance imaging (MRI). NMR research had increased during the 1940s and 1950s, and in 1959, a researcher used NMR with mice without causing harm, which facilitated the development and advancement of magnetic resonance imaging (MRI) for use with human subjects. However, MRI scanners were not common in hospitals until the mid-1980s (Kevles, 1997).

MRI technology uses the magnetic properties of brain tissue to generate a detailed image (Carter & Shieh, 2015). The images produced are three-dimensional and spatially precise (Lopes da Silva, 2010). For a full-body scan, a subject is positioned in a chamber where a magnetic field moves from the subject's lower extremities to their head (Carter & Sheih, 2015). For neurological research, only the head is scanned. Though a subject may experience some physical discomfort due to nausea or dizziness, the method is noninvasive, and image-capture is available from any angle. Another benefit is achieved by particular adjustment of the scanners, which can be varied to produce "images that highlight certain brain regions and provide contrast between specific kinds of neural tissue" (P. 11)

Figure 17. *An MRI Scanner (http://www.qhc.on.ca/mri-p760.php).*

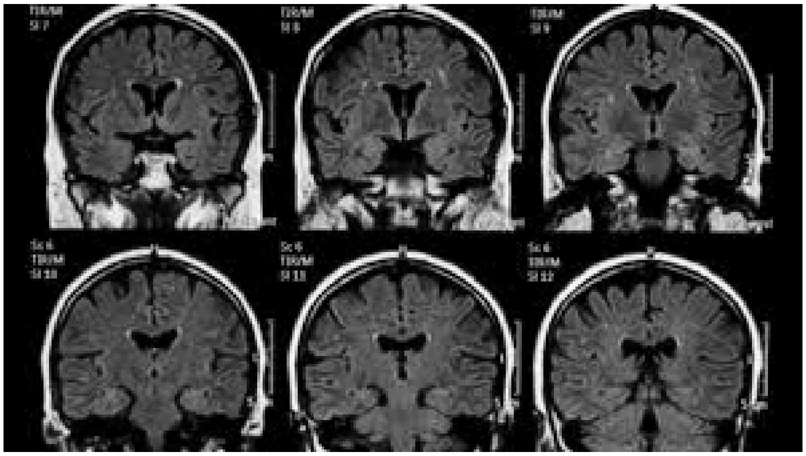


Figure 18. *An example of an MRI Scan (http://www.cbc.ca/news/technology/brain-sex-differences-1.3344954)*

Functional magnetic resonance imaging (fMRI) technology has almost entirely replaced the PET method. According to Kevles (1997), "Functional imaging, which had been PET's monopoly, is now possible with fMR. Functional MRI can track blood flow and the changes in oxygen within the bloodstream faster than PET and with at least ten times finer resolution" (P. 223). fMRI is similar to MRI technology but depends on the hemoglobin in the brain to determine changes in its oxygen levels. Images are taken before and after stimulus presentation or task performance. The images produced are four-dimensional figures with x, y, z coordinate planes, and time as the fourth dimension (Carter & Shieh, 2015). It is important to note that fMRI cannot determine causation, only correlation, though it is suitable for "longitudinal studies of the neural correlates of skill acquisition" (Stewart et al., 2003b, P. 204) as well as the exploration of the neural foundation of "cognition, emotion, sensation and behavior" (Carter & Shieh, 2015, P. 14). The benefits of fMRI are that the method is noninvasive and provides such specific data. However, the data can be difficult to detect and therefore requires many scans with the same stimuli, followed by statistical confirmation of the findings. There is a natural delay when oxygenated blood moves to a region of the brain and the resolution of the temporal data is inferior compared with other imaging methods. The fMRI machine emits a loud sound for the duration of the scan, which is a particular issue when studying the auditory system. The subject must remain still during scanning, but can complete tasks by responding with simple movements, such as with a keypress. Finally, this technology is expensive (Carter & Shieh, 2015).

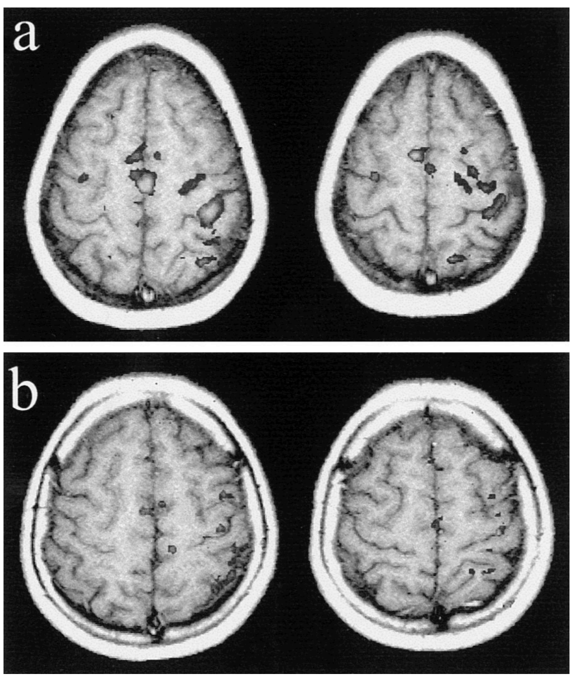


Figure 19. *An example of an fMRI scan (Krings et al., 2000, p. 191).*

5.1.4. Electroencephalography (EEG) and Magnetoencephalography (MEG)

Electroencephalography (EEG) is not a true imaging method but an effective measure for brain activity. According to Cantor (1999), EEG technology greatly advanced with the field of computer science in the 1960s and 1970s, though there are early accounts of electrical signals detected in humans and animals in the mid-to-late

19th-century. The first human EEG recordings were captured in 1924 by the psychiatrist Hans Berger (Boutros, 2011), and throughout the 1930s and 1940s, EEG was particularly beneficial for psychiatry and neurological science research (Cantor, 1999).

EEG technology is used on its own, or, more recently, with another imaging technique. This is an effective and beneficial approach to brain imaging (Jiang et al., 2014). “EEG combined with other imaging techniques, such as fMRI, can provide an excellent temporal and spatial representation of neural activity” (Carter & Shieh, 2015, P. 20). In this instance, a combination of methods broadens the potential for data collection, as researchers can know when and where brain events take place. EEG is also frequently combined with electrooculography (EOG). EOG is a method to measure eye movement and has been described in detail in the chapter concerning measurement by eye tracking. For cognitive neuroscience, the most useful measurement is of changes in the electric field, achieved by the placement of electrodes on the scalp (Comeau, 2016). When a stimulus is present, event-related potentials (ERPs) are measured (Dietrich & Kanso, 2010). Carter and Shieh (2015) explain that: “an ERP is a distinct, stereotyped waveform in the EEG that corresponds to a specific sensory, cognitive, or motor event...thus, neuroscientists can study the moment that the brain processes a stimulus with subsecond temporal resolution, allowing a noninvasive means of evaluating brain function in experimental subjects” (P. 20). As well, “qualitatively different processes are reflected by qualitatively different ERP waveforms” (Regnault, Bigand, & Besson, 2001, P. 243). The benefits of EEG are that it is noninvasive and one of the least expensive technologies.

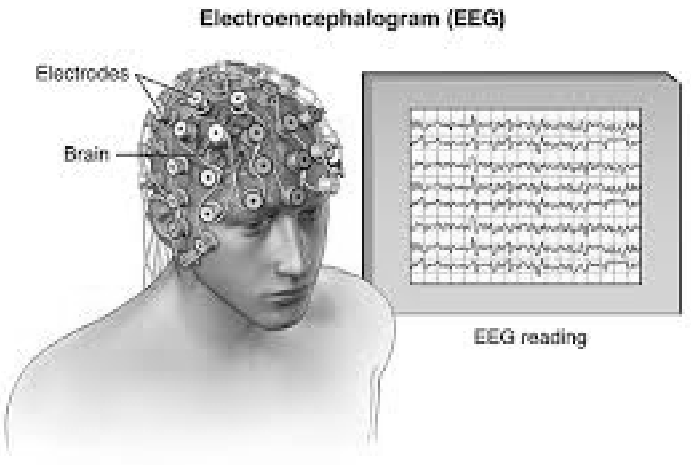
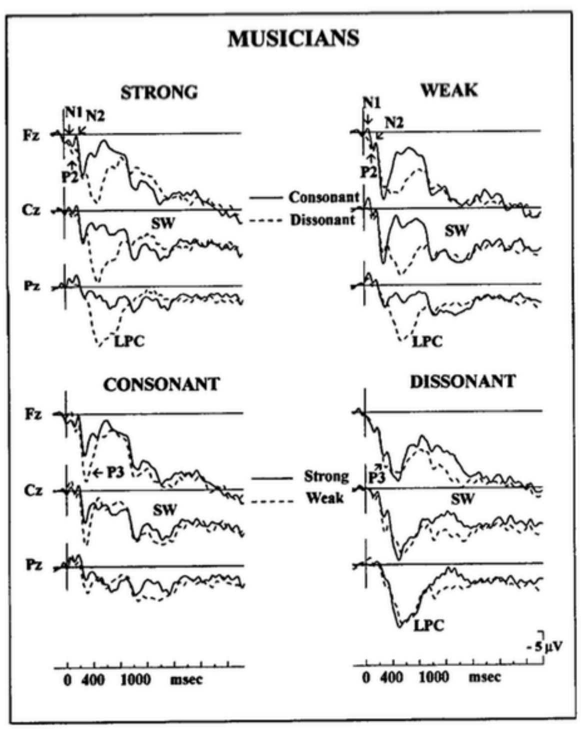


Figure 20. *A drawing of EEG electrodes placed on a subject, and of an EEG reading in progress (http://www.saintlukeshealthsystem.org/health-library/electroencephalogram-eeg).*



.

Figure 21. *An example of an EEG scan showing ERPs (Regnault et al., 2001, p. 245).*

Magnetoencephalography (MEG), a functional imaging technology, was first developed in the late 1960s (https://en.wikipedia.org/wiki/Magnetoencephalography). The technology measures the surface-level magnetic fields of the active brain (Lopes da Silva, 2010) and event-related fields (ERF), and requires fewer neurons than EEG to generate a signal for measurement (Carter & Shieh, 2015). However, it is these groupings of neurons in the cerebral cortex that are examined (Lopes da Silva, 2010). Schurmann, Raij, Fujiki, and Hari (2002) note the advantage of this technology is its "temporospatial precision" (P. 438). Its spatial resolution alone is not as strong as other methods; however, it combines well with fMRI imaging. The drawbacks of MEG are that it is difficult to determine the sources of the signals (Lopes da Silva, 2010) and it is a very expensive technology, in part because a specific facility is needed for operation (Carter & Shieh, 2015).

Figure 22. *An MEG scanner (https://www.britannica.com/science/autism/images-videos/An-autistic-child-viewing-a-movie-during-a-demonstration-of/155985)*

5.2. Music Reading Studies Using Neuroimaging Methods

Neuromusical research has employed imaging techniques. This includes the use of PET technology (Sergent, Zuck, Terriah, & McDonald, 1992) and MRI technology (Sergent, Zuck, Terriah, & McDonald, 1992). Presently, fMRI technology is frequently used for music research (Nakada, Fujii, Suzuki, & Kwee, 1998; Schön, Anton, Roth, & Besson, 2002; Stewart, Henson, Kampe, Walsh, Turner, & Frith, 2003a; 2003b; Wong & Gauthier, 2010). EEG technology has been used for music reading study (Stanzione, Grossi, & Roberto, 1990; Schön & Besson, 2002; 2005; Gunter, Schmidt, & Besson, 2003; Wong, 2010), as has MEG technology (Schurmann, Raij, Fujiki, & Hari, 2002). ERP measurement is especially valuable in neuromusical research. The record of cerebral electrical fluctuations provides a timeline of the processes activated with musical expectancy (Regnault, Bigand, & Besson, 2001).

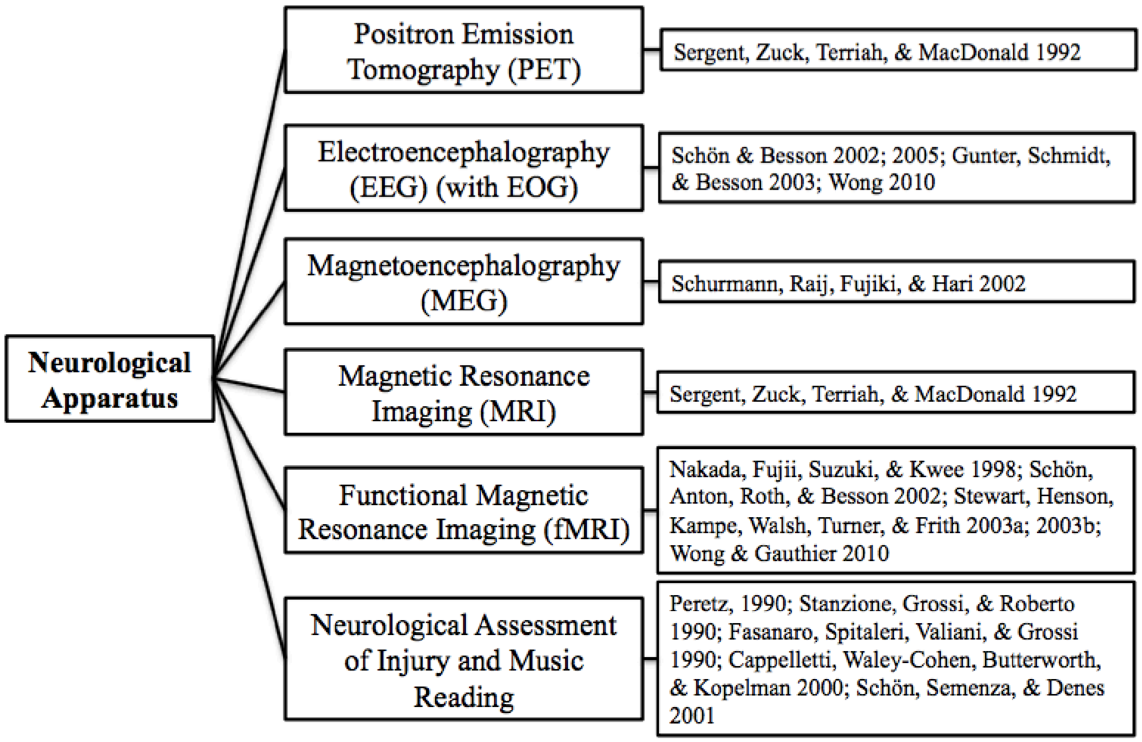


Figure 23. *Neurological measurement methods, subcategories, and studies.*

5.2.1. Positron Emission Tomography (PET)

Sergent, Zuck, Terriah, & MacDonald, 1992	
Type	Cognitive Process
Apparatus	PET
Production Method	Piano
Participants	10, professional pianists
Stimuli	J. S. Bach Partita
Control Group	No
Measurement	Functional Neuroanatomy

Positron Emission Tomography (PET) has been used for general music research, but rarely with the specific measure of music reading. However, Sergent, Zuck, Terriah, and MacDonald (1992) used PET, in combination with MRI, to study ten professional pianists as they completed reading, listening, and production tasks with a Bach Partita as the stimulus. PET was used to measure cerebral blood flow. The subjects were injected with a cerebral blood flow (CBF) tracer and measurement began 15 seconds afterwards, lasting for 60 seconds. The PET scanner generated three-dimensional images, which were aligned with the subsequent MRI images. Sergent and colleagues observed an activated network involved with the processing requirements of music reading and performance. The MRI portion of the experiment is explained in its respective section.

5.2.2. Magnetic Resonance Imaging (MRI) and Functional Magnetic Resonance Imaging (fMRI)

Sergent, Zuck, Terriah, & MacDonald, 1992	
Type	Cognitive Process
Apparatus	MRI
Production Method	Piano
Participants	10, professional pianists
Stimuli	J. S. Bach Partita
Control Group	No
Measurement	Functional Neuroanatomy

In addition to the use of PET technology, Sergent, Zuck, Terriah, and MacDonald (1992) used MRI methods to measure functional cerebral anatomy of ten professional pianists. The subjects' brains were imaged while participating in a variety of activation tasks: performing one voice of a Bach Partita with its respective hand, visually fixating on a screen, manually responding to a pattern of dots, reading a musical score with or without the accompanied performance, and listening to music. Sergent and colleagues observed an activated network involved with the processing requirements of music reading and performance, spanning the four lobes of the brain, as well as the cerebellum. The authors specifically noted that the superior parietal lobe is involved in the spatial features of the reading of notation.

Nakada, Fujii, Suzuki, & Kwee, 1998	
Type	Cognitive Process
Apparatus	fMRI
Production Method	Piano
Participants	8 pianists who spoke English and Japanese
Stimuli	Image, text, musical score
Control Group	No
Measurement	Neuroanatomy

Nakada, Fujii, Suzuki, and Kwee (1998) used fMRI technology to examine the cortical regions involved with music literacy, studying eight pianists who were proficient in English and Japanese. They compared the cortical activation of text presented in the subjects' languages with cortical activation of music notation. The subjects displayed activation in auditory and visual cortices, with an expertise effect in the right occipital cortex, which was only activated when reading a musical score.

Schön, Anton, Roth, & Besson, 2002	
Type	Cognitive Process
Apparatus	fMRI
Production Method	5-key keyboard
Participants	9, ages 24-50 years, with a minimum of 12 years piano experience
Stimuli	Music, verbal notation, Arabic numbers
Control Group	No
Measurement	Neuroanatomy

Similar stimuli had been previously used by Schön, Anton, Roth, and Besson (2002). They presented musical, verbal, or Arabic notation to pianists, who responded to the stimuli by playing a keyboard with their right hand or by using a keypress, respectively. Using fMRI, the authors found a similar activation pattern with all stimuli: the bilateral parietal lobes, the left sensorimotor cortex, and the right cerebellum. They were able to establish the use of the right parietal areas for music notation, specifically, and as with Nakada and colleagues (1998), that the right occipital cortex is significant for music reading.

Stewart, Henson, Kampe, Walsh, Turner, & Frith, 2003a; 2003b	
Type	Instructional
Apparatus	fMRI
Production Method	Piano
Participants	24, 2 groups of 12 subjects, with no reading/playing experience.
Stimuli	Simple RH melodies
Pre- and Post-Tests	No
Control Group	Yes
Measurement	Neuroanatomy; Cognition; Behaviour
Duration	15 weeks

Stewart, Henson, Kampe, Walsh, Turner, Frith (2003a; 2003b) used fMRI to measure the brain for functional modifications before and after the acquisition of music reading skill. The subjects did not have previous training in music reading. The authors found an explicit training effect in the superior parietal cortex when the subjects exercised their new ability to read music. This structure is located in the dorsal visual processing stream, an area significant for perceiving and coding the spatial features of visual stimuli. The authors state, "sight-reading for keyboard performance falls within the class of behaviors that the dorsal stream is known to subserve" (p. 206). The notation provides the position, which in turn directs the choice of appropriate piano key. Stewart and colleagues suggest that the activation of the superior parietal cortex is due to the "visuospatial translational element of music reading" (p. 81) and remark that the brain may be activated regardless of one's degree of proficiency.

Wong & Gauthier, 2010	
Type	Cognitive Process
Apparatus	fMRI
Production Method	Keypress
Participants	20, 10 music-reading experts and 10 novices
Stimuli	Objects, letters, mathematical symbols, single musical notes
Control Group	No
Measurement	Neuroanatomy; Perceptual Fluency

Wong and Gauthier (2010) could not replicate the expertise effect achieved by Nakada and colleagues (1998). Their 2010 fMRI study examined proficient music readers alongside novice musicians with little music reading experience. The subjects did not perform on an instrument but viewed single and sequential musical notes, mathematical symbols, or letters, with a keypress response. Wong and Gauthier also found involvement in a large neural network: the visual, auditory, audiovisual, somatosensory, motor, parietal, and frontal areas of the brain were activated while reading musical notation, either singly or sequentially. There was a higher selectivity for expert musicians, and the authors noted effects of expertise in the inferior temporal areas, among other parts of the brain.

5.2.3. Electroencephalography (EEG) and Magnetoencephalography (MEG)

Schön & Besson, 2002	
Type	Cognitive Process
Apparatus	EEG (with EOG)
Production Method	None
Participants	15, musicians
Stimuli	Clef with key and time signatures, then target note
Control Group	No
Measurement	ERPs; Reaction Times; Horizontal and Vertical EOG

Schön and Besson (2002) used EEG combined with electrooculography (EOG) to study the relationship between the visual representations of pitch and time in music reading. They measured the behavior and event-related potentials of eighteen musicians, and the horizontal and vertical EOG, respectively, while subjects completed notational matching tasks. In the second block, subjects used their right hand to perform each stimulus, which comprised the treble clef, a key and/or time signature, and a note with a given duration. Thus, in the second part of the experiment, the subjects processed pitch and temporal information simultaneously. Schön and Besson found that reaction times were shorter for pitch than for duration and for congruous target notes. They explain the results of the former:

> "First, in written music, pitch changes more frequently than duration. Thus, in a series of notes, the probability that two following notes change in pitch is higher than they change in duration. This is especially true in western music in which harmonic patterns are more salient than rhythmic patterns. Second, in reading/playing a score, a musician needs to anticipate the correct pitch before the onset of the note. In contrast, duration is determined by the off- set of the note, thus it is second in temporal priority. Third, pitch identification might be more language mediated than duration. Indeed, the pitch of notes can easily be named (A, B, etc.), while naming duration is not so frequent and it may also require compound words (dotted half note, etc.). Thus it may be that symbolic language mediation facilitated pitch identification and matching judgment" (p. 876).

Finally, the authors found no effect with the ERPs for the pitch element when measuring duration, and for the duration element when measuring pitch, which suggests independent processing for pitch and duration.

Schön & Besson, 2005	
Type	Cognitive Process
Apparatus	EEG (with EOG)
Production Method	None
Participants	14, musicians
Stimuli	Pairs of visual/auditory series of notes
Control Group	No
Measurement	ERPs; Reaction Times, Horizontal and Vertical EOG

Schön and Besson (2005) continued to explore musical notation, with the emphasis on one's auditory representation of the written score. The purpose of the experiment was to determine if expectancy can be based on visual stimuli only. Schön and Besson used EEG, again in combination with EOG, to measure ERPs, behavioural reaction times and errors, and horizontal and vertical EOG, respectively. There were both visual and auditory stimuli, where the auditory stimuli either matched or did not match the visual stimuli. The mismatched stimuli were either 'likely' or 'unlikely', thus creating a total of six conditions. The subjects were to judge the stimuli as quickly as possible and Schön and Besson found that judgment was faster and more accurate with matching conditions and with 'likely' stimuli. ERP data suggested that there are different processing mechanisms for likely and unlikely matching stimuli. The authors conclude that musicians can make some expectations based on the visual score; in this case, subjects were more consistent with 'likely' endings to stimuli. They write: "Thus, strong interactions seem to exist between visual and auditory musical codes, in that the representations built from visual stimuli influence the way auditory musical sequences are perceived" (p. 702).

Gunter, Schmidt, & Besson, 2003	
Type	Cognitive Process
Apparatus	EEG (with EOG)
Production Method	None
Participants	8 (exp. 1), from advanced music school with 17-26 years of classical Western music training; 11 (exp. 2), mean age 22.5 years, at advanced music school with 13-25 years of training in classical Western music.
Stimuli	Music phrases, and corresponding phrases with diatonic violation
Control Group	No
Measurement	Reaction Times (exp. 1); Bipolar Horizontal and Vertical EOG (exp. 2)

EEG, in combination with EOG, was also used by Besson and colleagues – (Gunter, Schmidt, & Besson, 2003) – in two other experiments: a behavioural measure of self-paced music reading, and an ERP measure of music reading. The purpose was to better realize the foundation of cognition in music reading, and to compare reading and auditory responses to "diatonic violations" (p. 744). The subjects were experienced with Western music and viewed a score measure-by-measure, determining the speed at which they would read. Gunter and colleagues found that measures containing violations resulted in slower reading, but that generally, the speed at which subjects read increased throughout the reading of the material. For the second experiment, along with the EEG recording, the authors measured bipolar horizontal EOG and bipolar vertical EOG to identify and characterize eye-movements. Again, subjects read a score that required two responses following the reading: if there was a violation present and if the subject recognized the music. Gunter and colleagues found that ERP data showed early and late effects for measures containing violations. The early effect is similar to auditory processing. The authors write: "Thus, it seems that violation detection in music relies on modality-independent and abstract processing mechanisms" (p. 750). The later effects, which occur after approximately 500 milliseconds, show greater differences between reading music and hearing music. The authors conclude that, with regard to ERP patterns, reading music is not the same as hearing music. Conversely, with text reading and speech, there are more similarities. These differences could be the result of distinct representations; text compared with notation, or the result of different cognitive processes for music and language.

Wong, 2010	
Type	Cognitive Process
Apparatus	EEG (with EOG)
Production Method	Keypress
Participants	22 (exp. 1, 2), experts with at least 10 years of music reading experience, novices who could not read music.
Stimuli	Single musical notes, Roman letters, or Pseudo-letters
Control Group	No
Measurement	Horizontal EOG; Perceptual Fluency (exp. 1); Perceptual Fluency; Holistic Processing (exp. 2)

Wong (2010) compared skilled musicians and novice musicians, with measures of perceptual fluency, expertise effects, EEG, and horizontal EOG. Subjects viewed a variety of stimuli, either singly or in sequence: musical notes, Roman letters, or pseudo-letters. Select stimuli were placed on a 5-line staff. Subjects were to quickly indicate, by keypress, when a stimulus was repeated. Wong found that skilled musicians had greater perceptual fluency with notational sequences than text sequences, and experienced less crowding in notation.

Schurmann, Raij, Fujiki, & Hari, 2002	
Type	Cognitive Process
Apparatus	MEG
Production Method	None
Participants	10, experienced professional-level musicians
Stimuli	Visual control, Auditory control, Note-sound associations, Auditory imagery, Control condition
Control Group	No
Measurement	Neuroactivation

A study using MEG technology by Schurmann, Raij, Fujiki, & Hari (2002) required professional musicians to perform an auditory imagery task where they analyzed the spatial position of musical notes as relating to the lines of the staff. From previous studies, the authors understood that the precuneus, an area of the brain between both the cerebral hemispheres and near the somatosensory cortex and visual cortex, was associated with visuospatial and visuomotor processing and visual imagery (Andersen et al., 1997, and Fletcher et al., 1995; Salenius et al., 1995, respectively). Schurmann and colleagues found neural activation in this area when viewing the musical notation, indicating a clear temporal sequence with the musician subjects. They also observed activation in a cortical network that included auditory association areas concerned with forming and remembering well-established audiovisual associations, specifically in the left and right occipital, left temporal auditory association, left and right premotor, left sensorimotor, and right inferotemporal visual association region.

5.2.4. Neurological Assessment of Injury and Music Reading

One approach of neurological music research is to study patients with brain damage. Studies of subjects with brain injury are common in literature. They include case studies of individual patients (Stanzione, Grossi, & Roberto, 1990; Fasanaro, Spitaleri, Valiani, & Grossi, 1990; Cappelletti, Waley-Cohen, Butterworth, & Kopelman, 2000; Schön, Semenza, & Denes, 2001) or group studies where all subjects have had the same injury (Peretz, 1990). This method of exploration is subject-dependent, both in the number of subjects and the area of study. However, it allows for comparison of healthy brain function with impaired brain function and provides new information with which to better understand the neural patterns involved in music cognition and production.

Peretz, 1990	
Type	Cognitive Process
Apparatus	None
Production Method	None
Participants	40, 20 unilateral brain-damaged patients and 20 normal controls
Stimuli	2-phrase musical excerpts; Rhythm excerpts
Control Group	Yes
Measurement	Music Processing

With the group of subjects in Peretz' 1990 study, all of whom had cerebrovascular lesions, she writes, "studying musical processing in brain-damaged nonmusicians provides important insights about the way the cerebral hemispheres cohere to provide a musical interpretation of the input" (p. 1203). According to Matthews (2008): "There is a growing body of literature on patients with focal lesions related to deficits in music perception that is complemented by new approaches using novel functional neuroimaging and electrophysiological paradigms to investigate music perception" (p. 461). However, when the objective is to provide further comprehension of music reading processes, case studies differ in methods from other neurological research in that they do not always employ technological apparatus, but often measure behavior with the use of a test battery to diagnose the patient and to learn about their musical strengths and limitations.

Stanzione, Grossi, & Roberto, 1990	
Type	Cognitive Process
Apparatus	None
Production Method	Verbal Naming; Transcription
Participants	1, 26-year old patient; Italian-speaking music teacher who suffered a head injury
Stimuli	Words; Nonwords; Musical Notes; Musical Symbols, Musical Score
Control Group	N/A
Measurement	Music Reading Ability

Two of these case studies consider means for notational processing. Stanzione, Grossi, and Roberto (1990) studied a patient whose diagnosis had been confirmed through electroencephalogram (EEG): damage in the left temporoparietal areas of the brain, including an intracerebral hematoma. This affected the patient's ability to read text and musical notation. Stanzione and colleagues employed a test battery to assess general neurological health and text reading. Measures of musical ability focused on reading time, and included the subject identifying single notes, successions of notes, time signatures, and musical symbols. The subject also completed dictation and transposition tasks. Additional subjects – all who were musicians - participated in the note reading tasks as controls. The authors found disruption only when the patient read increasing successions of notes, which allowed Stanzione and colleagues to propose a music-reading model comparable to text reading, that include both global and step-by-step reading processes. The authors suggest that the control subjects were able to group items into units, employing the global process, which resulted in faster completion of the task. The patient could no longer use this process and relied on the step-by-step, or in this case note-by-note, route.

Fasanaro, Spitaleri, Valiani, & Grossi, 1990	
Type	Cognitive Process
Apparatus	None
Production Method	Voice, Violin, Viola, Piano
Participants	1, 72-year-old patient; professional violinist. He suddenly developed loss of strength and hypoesthesia.
Stimuli	Words; Nonwords; Single Musical Notes; Note Sequences; Rhythms; Musical Symbols
Control Group	N/A
Measurement	Music Reading Ability

Fasanaro, Spitaleri, Valiani, and Grossi (1990) studied a professional musician who had alexia, which affects the ability to read, and amusia, or a disruption in the processing of pitch. They provided the patient with general neurological tests and text reading tasks. Measurement of musical items included the degree of receptive and expressive amusia, music reading and writing, and notational matching tasks. Fasanaro and colleagues found that the patient had impairment in both processing routes, and thus had to make adjustments in order to obtain an internal representation of the stimuli.

Cappelletti, Waley-Cohen, Butterworth, & Kopelman, 2000	
Type	Cognitive Process
Apparatus	None
Production Method	Piano, Verbal; Transcription
Participants	1, 51-year-old professional musician. At age 45, suffered episode of acute encephalitis with coma for 2 days.
Stimuli	Musical notes; Letters, words, numerals; Musical symbols; Other symbols; Auditory melodies
Control Group	N/A
Measurement	Music Reading Ability

Two case studies explore domain-specific impairments. The test battery used by Cappelletti, Waley-Cohen, Butterworth, and Kopelman (2000) began with a general neurological assessment and text and symbol-reading tasks. The patient read notes and musical symbols to measure music reading ability. The authors also measured general musical ability with tests of visual perception, auditory and visual melody recognition, transcoding of notation, music writing, music recall, understanding notational meaning, and music performance. They found that the patient had difficulty reading musical stimuli, and noted that the patient's injury were lesions at the right occipito-temporal junction and in the left postero-lateral temporal cortex. An earlier neurological study by Sergent and colleagues (1992) proposed a distinct neural network for the processing of music notation, namely in the superior parietal lobe. Cappelletti and colleagues suggest that the music-specific region identified by Sergent and colleagues is the same region affected by the patient's lesions, while the text-specific region was spared.

Schön, Semenza, & Denes, 2001	
Type	Cognitive Process
Apparatus	None
Production Method	Verbal; Transcription; Piano
Participants	1, 65-year-old female professional organist with a cerebral ischemic lesion on the left parieto-temporal region. Also, for experimental tests, 3 professional musicians (average age of 49) with 30+ years of experience who performed at ceiling.
Stimuli	Notes; Note Sequences; Musical Symbols
Control Group	N/A
Measurement	Music Reading Ability

Schön, Semenza, and Denes (2001) studied an organist with an ischemic lesion, which causes hypoxia due to lack of blood flow, on the left parieto-temporal region. This injury resulted in a disturbance of rhythm with melodic discrimination preserved. As well, the patient could not perform from memory. To begin, the authors measured general neurological health, and provided a sentence completion task. Musical measurement comprised transcoding processes and availability of symbol and notational knowledge and memory of these items. The transcoding processes investigated were notation to motor performance, verbal to notational, and verbal auditory to motor performance. The patient was able to function with language tasks when these were enveloped in context and syntax. Therefore, the patient's inability to read musical notation, especially in the bass clef, is identified as a domain-specific issue.

5.3. Summary

With **neuroimaging**, music reading tasks activate a network spanning the cortical lobes and cerebellum (Sergent, Zuck, Terriah, & McDonald, 1992), while Schurmann, Raij, Fujiki, and Hari (2002) suggest that auditory imagery creates an intricate network of temporospatial activations throughout several cortices: the left and right occipital, midline parietal (specifically the precuneus), left temporal auditory association, left and right premotor, left sensorimotor, and finally the right inferotemporal visual association region. fMRI studies of music reading offer images of activation in the auditory and visual cortices due to text and music reading (Nakada, Fujii, Suzuki, & Kwee, 1998), and general activation from text and music reading among the bilateral parietal lobes, left sensorimotor cortex, and right cerebellum, with the right parietal region involved in the reading of notation (Schön, Anton, Roth & Besson, 2002). There is also activation of the auditory, audiovisual, somatosensory, motor, parietal, and frontal regions with music notation (Wong & Gauthier, 2010).

Data from neuroimaging while reading music supports the position that pitch and duration are processed independently in music reading (Schön & Besson, 2002), that visual notation can aid with auditory comprehension (Schön & Besson, 2005), and that detecting incongruences in visual notation may involve working memory and a combination of integration and interpretation processes (Gunter, Schmidt, & Besson, 2003).

There are **expertise effects** with notes on or off the staff (N170 ERP effects for on-staff notes, and C1 and CNV ERP effects for off-staff notes), less crowding in musical notation, and effects expected with perceptual fluency (Wong, 2010). There is an effect from musical notation viewed in the right occipital cortex (Nakada, Fujii, Suzuki, & Kwee, 1998), and a training effect in the superior parietal cortex with learned music reading skill (Stewart, Henson, Kampe, Walsh, Turner, & Frith, 2003a; 2003b). Wong and Gauthier (2010) find expertise effects in inferior temporal areas and in the early visual cortex for notation.

The **study of brain injury** has offered a music-reading model consisting of global and step-by-step reading processes (Stanzione, Grossi, & Roberto, 1990), and found that impairment in one or both processes can be overcome by individual adaptation in order to understand musical stimuli (Fasanaro, Spitaleri, Valiani, & Grossi, 1990). Regions of the brain used for music processing may be distinctive from the areas used for text processing (Cappelletti, Waley-Cohen, Butterworth, & Kopelman, 2000), and domain-specific limitations may no longer be restrictive with more notational context and syntax (Schön, Semenza, & Denes, 2001). Peretz (1990) provides support for the belief that both hemispheres are needed for effective pitch sequencing, and that the left hemisphere of the brain is concerned with melodic details, while the right hemisphere is concerned with "global melody representations" (P. 1198).

Table 17. *Synthesis of summary tables for neurological apparatus and brain injury study.*

	Apparatus	Production Method	Participants	Measurement
Sergent, Zuck, Terriah, & MacDonald 1992	PET	Piano	10 (proficient)	Functional Neuroanatomy
Sergent, Zuck, Terriah, & MacDonald 1992	MRI	Piano	10 (proficient)	Functional Neuroanatomy
Nakada, Fujii, Suzuki, & Kwee 1998	fMRI	Piano	8 (proficient)	Neuroanatomy
Schön, Anton, Roth, & Besson 2002	fMRI	5-Key Keyboard	9 (proficient)	Neuroanatomy
Stewart, Henson, Kampe, Walsh, Turner, & Frith 2003a; 2003b	fMRI	Piano	24 (nonmusicians)	Neuroanatomy; Cognition; Behaviour
Wong & Gauthier 2010	fMRI	Keypress	20 (proficient; less proficient)	Neuroanatomy; Perceptual Fluency
Schön & Besson 2002	EEG (w EOG)	None	15 (proficient)	ERPs; RTs; Horizontal EOG; Vertical EOG
Schön & Besson 2005	EEG (w EOG)	None	14 (proficient)	ERPs; RTs; Horizontal EOG; Vertical EOG
Gunter, Schmidt, & Besson 2003	EEG (w EOG)	None	8; 11 (proficient)	RTs (exp. 1); Bipolar Horizontal and Vertical EOG
Wong 2010	EEG (w EOG)	Keypress	22 (proficient; nonproficient)	Horizontal EOG; Perceptual Fluency (exp. 1); Perceutal Fluency; Holistic Processing
Schurmann, Raij, Fujiki, & Hari 2002	MEG	None	10 (proficient)	Neuroactivation
Peretz 1990	None	None	40 (brain injury; controls)	Music Processing
Stanzione, Grossi, & Roberto 1990	None	Transcription	1 (26 yrs, brain injury)	Music Reading Ability
Fasanaro, Spitaleri, Valiani, & Grossi 1990	None	Voice, Violin, Viola, Piano	1 (72 yrs, brain injury)	Music Reading Ability
Cappelletti, Waley-Cohen, Butterworth, & Kopelman 2000	None	Piano; Transcription	1 (51 yrs, brain injury)	Music Reading Ability
Schön, Semena, & Denes 2001	None	Transcription; Piano	1 (65 yrs, brain injury)	Music Reading Ability

Table 18. *Summary of number of studies by classification.*

Measurement	**Cognitive Process**	**Instructional**	**Correlational**	**Test Validation**
Number of Studies	**13**	**1**	**0**	**0**

Conclusion

We identified the need for organization and stability in the research field of music reading assessment. We conducted a study based on systematic review of the literature concerned with music reading, in order to obtain the measurement methods used in assessing this skill. We identified the measurement methods and classified the studies by one of three broad categories: test measurements, eye-tracking measurements, or neurological measurements. Each broad category and subsequent subcategories were introduced with a brief history and descriptions of the apparatus used, with summaries of the relevant studies following.

When selected details of the body of studies are compared within a specific measurement category, there are differences between the apparatus used, with the type of test among the test measurements, or the type of technology employed. Within the test measurement category, there are few studies with a large number of participants. The exceptions to this are the studies of Zhukov (2014a) with 100 pianists, Boyle (1968) and Palmer (1976) with over 100 rhythm-reading participants, Demorest (1998) with 306 sight-singing participants, Miller (1988), McPherson (1994; 1995) and McPherson and colleagues (1997) with over 100 instrumentalists, and Gutsch (1966) with 771 rhythm-reading participants for a test validation study. Another finding is that most of the test measurements use author-composed material. This may indicate a need for standardized stimuli, such as the stimuli presented by the Watkins Farnum Performance Scale (Watkins & Farnum, 1954).

When the broad assessment methods are compared, we find similarities between the eye-tracking and neurological measurement categories. The experimental studies use a smaller number of participants and few child participants, with the exception of Bean (1938). There also appear to be more keyboard studies in this area, perhaps due to the ease of apparatus operation when a subject is seated at a keyboard and does not move their head in the same way as other instruments require.

Current research appears to be focused on cognitive processing, particularly because of access to eye-tracking and neurological apparatus, which is always improving. Conversely, correlational and instructional study has diminished, though these areas may further pedagogical and psychological fields and offer more options for assessment to the piano teacher or classroom music instructor. All measurement methods provide very few validation studies, if any.

Conducting a survey based on systematic review proved to be a successful method of identifying and gathering relevant information. The search process was fruitful, even during detailed examination of a study, which often provided direction for obtaining additional studies. Identifying the information needed from each study facilitated readable summary charts of each study, and a means for comparison. This comparison made clear a lack of studies in a specific area, inconsistencies in the literature, and trends in this field of research. As well, the classification of music reading research provides organization, offering a clearer understanding of how the methods have developed and where research in this field is directed. Researchers, pedagogues, and music teachers alike may access this classification in order to guide their assessment of music reading ability.

Table 19. *Complete summary of number of studies by classification.*

	Cognitive Process	**Instructional**	**Correlational**	**Test Validation**
Test Measurements	**17**	**14**	**18**	**3**
Eye-Tracking Measurements	**26**	**1**	**0**	**0**
Neurological Measurements	**13**	**1**	**0**	**0**

Test Measurements: Limitation of Available Studies and Need for Future Research

According to Pagan (1970), the development of music reading skills, particularly to read with precision and speed, is "widely recognized by educational institutions" (p. 2). Pagan notes that university-level curriculum focuses on reading and hearing music accurately. Kornicke (1992) writes of great interest in sight reading, supported by the fact that many research and pedagogical sources examine the skill. In the results of a questionnaire by Zhukov (2014b), pianists ranging from high school- to university-age and from private studios and institutions collectively stated that sight reading was 'important' or 'very important'.

Assessment of music reading and sight reading results in a better understanding of the processes involved. This, in turn, helps to continue instruction and practice in developing reading abilities. However, Kornicke (1992) writes, "Given the recognized importance of sight-reading, it is surprising to note considerable limitations in the sight-reading literature among pedagogical, music education, and psychological sources" (p. 3). Regarding music education, Scripp (1995) remarks, "despite intensive instruction...music students show a lack of coordination of music reading skills across various dimensions of production and perception. That is, while reading skills may be considered essential tools for developing understanding of the literature, the course of their development in formal education is uncertain, at best, and their coordination across modes of performance is extremely limited" (p. 34). Thus, there are limitations to the field of music reading, and the body of music reading research literature.

To begin, there are uncertain definitions regarding aspects of music reading assessment. The term 'sight reading' is incongruent with its scoring measures (Kornicke, 1992). Kornicke writes that the use of dynamics and *rubato* is important to proficient sight readers to portray the quality of a piece, though most scoring methods instead focus on pitch and rhythm. Subtests of test batteries that have been purposed to measure one specific skill have at times inadvertently measured many skills (Law & Zentner, 2012). Pagan (1970) states, "the word 'measure' has a multiplicity of meanings and applications...its precise meaning is dependent upon several factors: the observer, the object, the purpose of measurement...most people expect that a measurement will be expressed by a quantity" (p. 8-9). A common measure in music reading research is the comparison of musicians with nonmusicians. However, Law and Zentner (2012) remark that this does not mean the nonmusician subject does not possess musical ability, but that it may yet develop, or its development has been inhibited. Finally, there is no consensus on the best measurement of musical ability with "objective tasks" (p. 2), partly due to the complex meaning of music and musical ability.

With inconsistency in measurement methods, the resulting reading research is disjointed (Zhukov, 2014b). There are a variety of assessment approaches, including the use of tests and scales or error measurement, and scoring by expert examiners or the use of technology (Lemay, 2008). As early as Watkins' (1942) study, he wrote "in music...after ability tests have been carefully constructed, they have been compared with crude measure of students' achievement" (p. 4). As well, Pagan (1970) writes that as research provides new information, there are only a few experts, specifically psychometrists, who can take the information from insulated studies and develop them "into cohesive statements of principles of psychological measurement" (p. 7).

The items of assessment within measurement also vary. Scripp (1995) writes that music reading research is "scattered in focus and methodology" (p. 3), particularly because studies present a range of samples, design, and analysis. Test batteries have employed differing amounts of stimuli, and the stimuli itself might vary in its performance duration. Auditory stimuli, if recorded with musicians, may have errors or irregularities in timing or sound quality (Law & Zentner, 2012). Stimuli for piano studies have often consisted of one line of music instead of two, and Kornicke (1992) writes, "these studies have sacrificed a crucial factor in keyboard sight-reading" (p. 8). In agreement, Betts and Cassidy (2000) write, "the reading of multiple notes on two staves, frequently in different clefs, along with the independent involvement of both hands and often pedal, requires different types of skills than those defined in research on single-line reading" (p. 152). Betts and Cassidy add that generalizing results from single-melody stimuli degrades piano reading research. Scripp (1995) remarks that scorer judgment is challenged with vocal or instrumental production, where a pitch can change gradually. In the same way, tempo can change subtly during performance. Novice readers are often inconsistent with the type of errors committed, as well as the regularity of errors, and this poses a challenge for evaluation (Gudmundsdottir, 2003). According to Law and Zentner (2012), measurement achieved with advanced apparatus, such as in neurological research, compared with a "dichotomous predictor of uncertain validity" (p. 2), such as musicians versus nonmusicians, results in inefficient and diminished data. There are different formats for responses, and little regulation over response bias or patterns of guesses by subjects (Law & Zentner, 2012). Zhukov (2014b) notes that studies concerning sight reading instruction present varying teaching methods and have not been replicated. For assessment of musical ability, Law and Zentner (2012) state that a lack of measurement apparatus results in research with self-reported data, which is then used to approximate musical ability. In the same manner, pedagogy literature for music reading offers instructional approaches, but this information is often relying on the experience and narratives of individual teachers, which limits its application potential (Kornicke, 1992). Finally, there are "significant differences among the testing procedures in terms of external support conditions" (Scripp, 1995, p. 103), as well as differences in error coding.

Music reading tests with various production methods have their own limitations. For vocal sight reading assessment, it is difficult to assess subjects individually in the usual rehearsal environment, individual testing and scoring require large amounts of time, scoring is often subjective, and there is "no equivalent test...universally accepted or used by researchers interested in vocal sight-reading" (Henry, 2001, p. 22). There are very few studies of rhythm reading assessment (Boyle, 1968). Finally, the evaluation of music reading readiness is difficult to implement in schools, and thus the small amount of subjects in the available literature diminishes validity (Legore, 1981).

In general, the reliability and validity of assessment is affected by discrepancies in assessment methods. “The procedures used for inferring test validity and reliability are tenuous by contemporary standards” (Law & Zentner, 2012, p. 3). Kornicke (1992) describes reading tests, including those composed by Eaton (1978) and Lowder (1973), as limited because of their lack of scrutiny for reliability or validity. According to Lemay (2008), the many approaches to assessment have not been compared. It is difficult to know if the assessment methods are equivalent, and subsequently, the reliability of the data suffers. Legore (1981) comments that “standardized music achievement tests are available but...lose reliability below fifth grade while research indicates that formation of concepts necessary for music reading readiness occurs at these lower grade levels” (p. 2).

For correlation studies, Mishra (2014) writes, “the research appears to have grown without organization...variables seem chosen because the data are convenient, rather than to test a hypothesis...the sheer number of variables makes it impossible to test (or retest) all possible variables within one study. This has led to widely disparate results with little to connect them” (p. 453). McPherson (1994) states that correlation studies often consist of small sample sizes and varying production methods. Oftentimes, the study is not directed to factors of sight reading and thus measurement methods or required tasks digress from the purported objectives of the study.

The limitations affect music reading education. Miller (1988) writes, “there remains substantial disagreement about how to teach music reading” (p. 1). Miller remarks that there are wide-ranging instructional approaches found in pedagogy and classroom literature. According to Zhukov (2005), “no studies have been replicated and few use similar approaches. This makes it difficult for teachers to find a reliable method for teaching sight-reading and forces them to rely on their own experience and materials.

There exists an urgent need for the development of effective means of teaching sight-reading and for appropriate teaching resources" (p. 2). An example is described by Zhukov (2014b), where the concept of the eye-hand span is difficult to convey in the private piano studio setting; a teacher encouraging a student to 'look onward' would not instinctively develop the eye-hand span. Scripp (1995) remarks that there is not a standard "theory of music reading development" (p. 3), which results in music educators examining the available research without a framework to guide them. Accordingly, Zhukov (2014b) asserts that a lack of guidance results in a deficiency of sight reading instruction.

The available literature on music reading may be insufficient to understand how music reading skill is acquired and developed (Scripp, 1995). However, research is continuing to provide useful information for further research and for teaching. For example, Kornicke (1992) found that aural imaging and ear training is important to sight reading, and states that the two components need to be separated for measurement and for instructional purposes. Zhukov (2014b) found that teaching approaches using accompanying experience, rhythm practice, and familiarity with musical styles show potential for reading development and success.

In 1970, Pagan wrote, "In spite of the progress made in the past twenty or twenty-five years, the science of psychological measurement is still in the process of development, and new ideas and alterations in presently accepted theories may be expected" (p. 8). This is still true today, and the field of music reading assessment continues to be developed and refined. According to researchers of collected studies, further research areas should include "a group measure of the ability of musicians to read the harmonic aspect of a musical score" (Pagan, 1970, p. 3), the development of an objective measure for music reading readiness, in order to help reduce the attrition rate of novice instrumentalists (Legore, 1981), and the application of the method used by Demorest (1998) of testing subjects individually to instrumental sight reading. He writes, "It would be useful to know just how much or how little testing is needed...and whether improvement would continue over a longer period of time" (p. 191). Zhukov (2006) encourages the development of an effectual program for sight reading, for "all stages of learning a musical instrument" (p. 5), and after collecting pianists' thoughts regarding sight reading, Zhukov (2014b) conveys the need for an inclusive training program that employs cooperative performance, rhythm training, and pattern recognition within musical styles. For correlative studies, McPherson (1995) encourages the study of rehearsal behaviour and the study of skill acquisition for performance abilities and Mishra (2014) states the need for further examination of the irregularity within correlation data, to "begin to explain the correlations" (p. 463). Mishra (2013) recommends examining the proposed idea that sight reading skill is able to be taught and developed.

Eye-Tracking Measurements: Limitation of Available Studies and Need for Future Research

There are certain limitations to eye-tracking research in the area of music reading, in part because the study of eye movement with music can be inconsistent and therefore confusing. The use of different apparatus makes it difficult to compare studies directly, and other areas of music research, such as cognition and perception studies have progressed with developments in neuroscience technology. Generally, eye-tracking apparatus impedes head movement, which is crucial for natural motion when singing or performing an instrument, and therefore limits the measurement of music reading (Madell & Hébert, 2008). Madell and Hébert also write:

> While each of these studies makes a valuable contribution to the field, the collective data are inconclusive and rather confusing. Many of the inconsistencies may be attributed to sheer lack of statistical analysis. Others are a result of methodological issues such as small participant pools and stimuli sets, shifting paradigms and tasks, and varying definitions of eye movement parameters (e.g., eye-hand span) and skill level. Most importantly, however, the lack of both focus on musical structure and convincing links between structure and eye movements makes comparisons across studies difficult and isolates music reading studies from other music cognition and perception studies (p. 166).

An area of greater achievement in research is with text reading. Compared with text reading research, there is a limited amount of inquiry of the perceptual span (Gilman & Underwood, 2003). Unlike written text, piano notation often ranges horizontally and vertically, especially with harmonic notation. Gilman (2000) writes, "hence for music reading, the perceptual span should be measured in terms of its height as well as its width. As yet, there have been no reports of experiments that have investigated musicians' perceptual spans while reading homophonic or polyphonic music" (p. 335). Concerning the eye-hand span, many studies have employed small numbers of subjects. The studies of Truitt and Colleagues (1997), and of Furneaux and Land (1999) each included eight subjects (Penttinen, Huovinen, & Ylitalo, 2015). For the study of parafoveal vision, it is difficult to deduce whether notation influences eye movement, or whether notation is processed, in whole or in part, near to fixation (Madell & Hébert, 2008). Furneaux and Land (1999) remark that few studies have examined eye movements with music reading, particularly with study of the eye-hand span. Penttinen, Huovinen, and Ylitalo (2012) write, “one of the potential strengths of the eye-tracking methodology is that it could help us understand the differences in the cognitive requirements of such different music-reading contexts” (p. 793). However, the authors describe one of the shortcomings of this field: an imprecise definition of ‘sight-reading’, particularly, which results in difficulty interpreting eye-movement data. “In recording eye movements with the accuracy of milliseconds, we need to be aware that familiarity with the notated music tends to affect the reading” (p. 793). In general, eye movement research of music reading has not yet reached the same level of complexity as with research of text reading, and Madell and Hébert (2008) suggest that music researchers continue to adapt and use successful models from text reading studies to better understand the cognitive foundation of music reading. Finally, Penttinen, Huovinen, and Ylitalo (2015) raise an interesting point: “at present, there may be no universally accepted principles to guide sight-reading pedagogy: approaches to teaching music reading vary from one instructor to another. Informal accounts nevertheless attest to some pedagogical approaches, which might become suspect in the light of eye-movement research” (p. 48). Regular and reliable music reading research with eye-tracking will also enrich piano pedagogy, and permit consistent teaching methods.

Neurological Measurements: Limitation of Available Studies and Need for Future Research

Schön, Anton, Roth, and Besson (2002) note that it remains unclear as to which regions of the brain are essential for the study of music reading. There are numerous studies of text reading compared with studies of music comprehension and production, with the limited number of neuropsychological studies of "music reading mostly concerned with the dissociation between reading the alphabetic code and reading musical notation" (Schön, Semenza, & Denes, 2001, p. 408). This is, in part, because many of the case studies are with a single subject, often with a lesion as the disruptor, and because limited studies have employed neuroimaging technology (Schön et al., 2002). Schön, Semenza, and Denes (2001) state that there is a shortage of professional musicians to study as patients, as these are the subjects who possess the required complex musical accomplishments. As well, the researchers need to be musically proficient. Music reading is often studied as a complete system, but can be studied in subdivisions of transcoding processes: visual to auditory, visual to motor, and visual to verbal. "Thus, while some cognitive operations and neural networks might be common to these three types of transcoding, others may well differ" (Schön, Anton, Roth, & Besson, 2002, p. 2285). According to Stewart and colleagues (2003b), pianists, both children and adults, are introduced to reading notation early in their training. However, few neuroimaging studies have child or adolescent subjects. With regard to the lack of child subjects, this could be on account of the design of the technology: scalp electrodes, injections, or placement in a chamber could be distressing (Hodges, 2015). However, Hodges also remarks that the field of neuroscience is providing a better understanding of what happens to the brain during formative musical training in childhood. Schön, Semenza, and Denes (2001) conclude: "Thus, information about the neural organization of music in the brain is still very fragmentary, and several basic abilities, such as music reading and writing, still remain to be investigated (p. 408).

Limitations of the Thesis and Future Study

Many studies on music reading were reviewed in this thesis. It appears that the categories of test, eye-tracking, and neurological measurements will also be suitable for continued classification of future research studies. However, available studies may have been overlooked due to a variety of reasons: 1. The studies did not match our predetermined dating; 2. The studies are unpublished and therefore difficult to locate; 3. The studies were not found within the search criteria; 4. The measurement method of a particular study was not included or explained clearly to allow us to classify the study. With the rate of technological advancement, future research may employ a new measurement method requiring an original category.

In the descriptions of the apparatus, particularly with eye-tracking and neurological apparatus, we have not gone into great detail about the biological processes of the brain or eye; instead providing an overview. The incredible detail and knowledge base needed, as well as being mindful of the objective to keep this document as a summary and reference resource, prevents us from providing an excessive amount of information. Thus, when wanting to learn more about the results of a study, one should reference the original study or learn more about the measurement areas. Likewise, the author may have misinterpreted results or omitted a piece of information when reporting on the various studies. A number of studies employ many areas of apparatus for measurement (Salis, 1977; 1980; Thompson, 1985; Goolsby, 1994a; Gilman, 2000). In these cases, the central measurement method was chosen in order to classify the study.

The study may be more appropriate in another category, depending on the lens used to interpret it. Finally, as stated in the introduction, the majority of the stimuli presented in these studies, in all categories, lend themselves almost entirely to Western tonal music. As well, the teaching methods are considered ‘traditional’, and the majority of subjects are familiar with an established instructional method. Though a few studies use 20th-century, or ‘contemporary’ repertoire (Bean, 1938; Micheletti, 1980; Gilman, 2000; Zhukov, 2014a), this generally could be considered a limitation in terms of the scope of repertoire and educational methods used.

However, in spite of the limitations, our study is able to identify the areas in need of novel of continued study. Further research could explore music perception studies (e.g. Bigand, 1997; Boltz, 1998; Waters & Underwood, 1999), which may prove to be a significant category for music reading assessment methods. Though our search was intensive, a broad compilation of music reading literature from a variety of musical cultures would introduce a diverse assemblage of repertoire styles and educational experiences and might introduce new methods of reading assessment. To strengthen the reliability and validity of the current studies, many could be replicated with larger sample sizes, if possible, while also addressing the limitations of the particular study. In general, the testing apparatus requires repetition of study to allow for true validation and reliability measures. For keyboard reading research, the methods specific to this area need to be supported by further and repeated study. The proven methods can then be adapted for the private piano studio to aid teachers in evaluating student’s music reading needs and achievement.

The study of cognitive processing will be a chief focus for future research, but opportunities are available to also develop instructional and correlational research, especially in combination with current apparatus and our ever-developing knowledge base. Growth in this area will provide opportunities for music pedagogues, researchers, and students alike in the pursuit of music reading comprehension, skill, and assessment.

References

Adèr, H. J., Mellenbergh, G. J., & Hand, D. J. (2008). *Advising on research methods: A consultant's companion*. Huizen, The Netherlands: Johannes van Kessel.

Agay, D. (1981). *Teaching piano: A comprehensive guide and reference book for the instructor* (Volume 1). New York, NY: Yorktown Music Press, Inc.

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

Aliferis, J. (1957). The Aliferis Music Achievement Test. *American Music Teacher, 7*(1), 6-7, 20.

Anon. (n. d.) *Positron emission tomography (PET).* Retrieved June 12, 2016, from http://bme.ucdavis.edu/cmgi/services/positron-emission-tomography-pet/

Anon. (2016). *Business units and fields of use*. Retrieved March 29, 2016, from http://www.tobii.com

Anon. (2016). *Dual Purkinje eyetrackers*. Retrieved June 11, 2016, from http://ppw.kuleuven.be/home/english/research/lep/resources/purkinje

Arden, G. B. (2006). Origin and significance of the electro-oculogram. In J. R. Heckenlively, G. B. Arden (Eds.), *Principles and practice of clinical electrophysiology of vision* (Second Ed). (pp. 123-138). Cambridge: The MIT Press.

Associated Board of the Royal Schools of Music. (2008). *Specimen sight-reading tests*. London: Associated Board of the Royal Schools of Music.

Baker-Jordan, M. (2004). *Practice piano pedagogy: The definitive text for piano teachers and pedagogy students*. Miami, Fla: Warner Bros.

Banton, L. J. (1995). The role of visual and auditory feedback during the sight-reading of Music. *Psychology of Music, 23*(1), 3-16.

Barea, R., Boquete, L., Rodriquez-Ascariz, J. M., Ortega, S., & López, E. (2011). Sensory system for implementing a human-computer interface based on electrooculography. *Sensors, 11*(1), 310-328.

Bastien, J. W. (1995). *How to teach piano successfully* (3rd ed.). San Diego, CA: N.A. Kjos Music Co.

Bastien, J. W., & J. S. (1985). *Bastien piano basics*. San Diego, CA: N. A. Kjos Music Co.

Bean, K. L. (1938). An experimental approach to the reading of music. *Psychological Monographs, 50*(6), i-80.

Bebeau, M. J. (1982). Effects of traditional and simplified methods of rhythm-reading instruction. *Journal of Research in Music Education, 30*, 107-119.

Bennett, E. & Capp, H. (1968). *Musical encounters: Sight reading and ear tests*. Oakville, Ont: F. Harris Music Company.

Benschop, R. (1998). What is a tachistoscope? Historical explanations of an instrument. *Science in Context, 11*(1), 23-50.

Berk, L. E. (2001). *Development through the lifespan* (2nd ed.). Needham Heights, MA: Allyn & Bacon.

Berlin, B. & Markow, A. (2002). *Four star sight reading and ear tests: Daily exercises for piano students*. Mississauga, Ont.: Frederick Harris Music.

Betts, S. L., & Cassidy, J. W. (2000). Development of harmonization and sight-reading skills among university class piano students. *Journal of Research in Music Education, 48*(2), 151-162.

Bigand, E. (1997). Perceiving musical stability: The effect of tonal structure, rhythm, and musical expertise. *Journal of Experimental Psychology: Human Perception* *and Performance, 23*(3), 808-822.

Boltz, M. G. (1998). The processing of temporal and nontemporal information in the remembering of event durations and musical structure. *Journal of Experimental Psychology: Human Perception and Performance, 24*(4), 1087-1104.

Boutros, N. (2011). Historical review of electroencephalography in psychiatry. In N. Boutros, S. Galderisi, O. Pogarell, & S. Riggio (Eds.), *Standard electroencephalography in clinical psychiatry: A practical handbook.* Chichester, West Sussex: Wiley-Blackwell.

Bowles, R. W. (1971). *Belwin-Mills singing achievement test*. Melville, NY: Belwin-Mills.

Boyle, J. D. (1968). *The effect of prescribed rhythmical movements on the ability to sight read music* (Order No. 6817359). Available from ProQuest Dissertations & Theses Global. (302315007). Retrieved from http://search.proquest.com/docview/302315007?accountid=14701

Boyle, J. D. (1973). Iowa Tests of Music Literacy by Edwin Gordon (REVIEW). *Journal of Research in Music Education, 21*(3), 281-283.

Bunge, S. A., & Kahn, I. (2009). Cognition: An overview of neuroimaging techniques. In L. R. Squire (Ed.), *Encyclopedia of Neuroscience* (Vol. 2. pp. 1063-1067). Retrieved from http://www.sciencedirect.com/science/referenceworks/9780080450469

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

Cantor, D. S. (1999). An Overview of Quantitative EEG and Its Applications to Neurofeedback. In A. Abarbanel & J. R. Evans (Eds.), *Introduction to quantitative EEG and neurofeedback* (pp. 3-27). Burlington, VT: Burlington Academic Press.

Carpenter, R. H. S. (1988). *Movements of the eyes* (2nd ed.). London: Pion Limited.

Carter, M., & Shieh, J. C. (2015). *Guide to research techniques in neuroscience, 2nd ed.* London: Academic Press.

Cappelletti, M., Waley-Cohen, H., Butterworth, B., & Kopelman, M. (2000). A selective loss of the ability to read and to write music. *Neurocase, 6*(4), 321-332.

CBC/Radio-Canada. (2016). *Brains of men and women aren't distinctly different, MRI study shows.* Retrieved June 12, 2016, from http://www.cbc.ca/news/technology/brain-sex-differences-1.3344954

Chadwick, J. E. (1933). Predicting success in sight-singing. *Journal of Applied Psychology, XVII*, 671-674.

Clark, F., Goss, L., & Holland, S. (2000). *The music tree: A plan for musical growth at the piano*. Miami, FL: Summy-Birchard Inc.

Colwell, R. (1969). *MAT Music Achievement Tests 1 and 2*. Chicago: Follett Educational Corporation.

Colwell, R. (1970). The development of the Music Achievement Test series. *Bulletin of the Council for Research in Music Education, 22*, 57-73.

Colwell, R. J. (1991). Evaluation. In R. J. Cowell (Ed.), *Basic concepts in music education II* (pp. 247-278). Niwot, CO: University Press of Colorado.

Colwell, R. (2002). Assessment's potential in music education. In R. Colwell, C. P.

Richardson (Eds.), *The new handbook of research in music teaching and learning* (pp. 1128-1158). New York: Oxford University Press.

Comeau, G. (2009*). Piano pedagogy: A research and information guide*. New York: Routledge.

Comeau, G. (2016). Methods for studying the brain [Powerpoint slides].

Coombs, D. S. (2012). An untrained eye: the tachistoscope and photographic vision in early experimental psychology. *History and Technology, 28*(1), 107-117.

Coupland, S. G. (2006). Electrodes for visual testing. In J. R. Heckenlively, G. B. Arden (Eds.), *Principles and practice of clinical electrophysiology of vision* (Second Ed). (pp. 245-254). Cambridge: The MIT Press.

D'Anselmo, A., Guiliani, F., Marzoli, D., Tommasi, L., & Brancucci, A. (2015). Perceptual and motor laterality effects in pianists during music sight-reading. *Neuropsychologia, 71*, 119-125.

Demorest, S. M. (1998). Improving sight-singing performance in the choral ensemble: The effect of individual testing. *Journal of Research in Music Education, 46*(2), 182-193.

Derdeyn, C. P. (2007). PET imaging in carotid stenosis. In B. Schaller (Ed.), *Imaging of Carotid Artery Stenosis* (pp. 85-102). New York: Springer Wien.

Dietrich, A., & Kanso, R. (2010). A review of EEG, ERP, and neuroimaging studies of creativity and insight. *Psychological Bulletin, 136*(5), 822-848.

Drai-Zerbib, V., & Baccino, M. T. (2005). L'expertise dans la lecture musicale: Integration intermodale*. L'année psychologique, 205*(3), 387-422.

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

Drake, C., & Palmer, C. (2000). Skill acquisition in music performance: Relations between planning and temporal control. *Cognition, 74*(1), 1-32.

Eastlund Gromko, J. (2004). Predictors of music sight-reading ability in high school wind players. *Journal of Research in Music Education, 52*(1), 6-15.

Eaton, J. L. (1978). A correlation study of keyboard sight-reading facility with previous training, note-reading, psychomotor, and memorization skills (Doctoral dissertation, Indiana University*). Dissertation Abstracts International*, A 39 (07), 4109.

Elliott, C. A. (1982a). The relationships among instrumental sight-reading ability and seven selected predictor variables. *Journal of Research in Music Education, 30*(1), 5-14.

Elliott, C. A. (1982b). The identification and classification of instrumental performance sight reading errors. *Journal of Band Research, 18*(1), 36-42.

Encyclopaedia Brittanica, Inc. (2016). *Media for autism*. Retrieved June 12, 2016, from https://www.britannica.com/science/autism/images-videos/An-autistic-child-viewing-a-movie-during-a-demonstration-of/155985

Erlings, B. (1977). A pilot investigation of relationships between elementary keyboard sightreading achievement by music majors in college and selected musical profile tests. *Bulletin of the Council for Research in Music Education*, 14-17.

Faber, N. & Faber, R. (1993*). Piano adventures: The basic piano method*. Ann Arbor, MI: Faber Piano Adventures.

Farnum, S. E. (1969). *The Farnum string scale.* Winona, MN: Hal Leonard.

Fasanaro, A. M., Spitaleri, D. L., Valiani, R., & Grozzi, D. (1990). Dissociation in musical reading: A musician affected by alexia without agraphia. *Music*

Perception, 7(3), 259- 272.

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.

Finn, C. & Morris, E. (1998). *Beanstalk's basics for piano*. Florence, Ky: Willis Music.

Fitzpatrick, D. & Mooney, R. D. (2012). Vision: The eye. In D. Purves, G. J. Augustine, D. Fitzpatrick, W. C. Hall, A-S. Lamantia, L. E. White (Eds.), *Neuroscience* (5th ed.). (pp. 229-256). Sunderland, Massachusetts: Sinauer Associates, Inc. Publishers.

Fitzpatrick, D. & Mooney, R. D. (2012). Central visual pathways. In D. Purves, G. J. Augustine, D. Fitzpatrick, W. C. Hall, A-S. Lamantia, L. E. White (Eds.), *Neuroscience* (5th ed.). (pp. 257-276). Sunderland, Massachusetts: Sinauer Associates, Inc. Publishers.

Flohr, J. W. & Hodges, D. (2002). Music and neuroscience. In R. Colwell, C. P. Richardson (Eds.), *The new handbook of research in music teaching and learning* (pp. 991-1008). New York: Oxford University Press.

Fourie, E. (2000). The role of visual sensory memory and chunking in the development of sight reading. In A. Snedden (Ed.), *Proceedings of the 4th Australian National Piano Pedagogy Conference. Embracing the New Millennium* (pp. 25-34). The University of Western Australia, Nedlands: Callaway International Resource Centre for Music and Music Education.

Fourie, E. (2004). The processing of music notation: Some implications for piano sight-reading. *Journal of the Musical Arts in Africa, 1*(1), 1-23.

Furneaux, S. & Land, M. F. (1999). The effects of skill on the eye-hand span during musical sight-reading. *Biological Sciences, 266*(1436), 2435-2440.

Geisinger, K. F. (2013). Reliability. In K. F. Geisinger, B. A. Bracken, J. F. Carlson, J-I. C. Hansen, N. R. Kuncel, S. P. Reise, M. C. Rodriguez (Eds*.), The APA handbook Of testing and assessment in psychology, Vol. 1: Test theory and testing and assessment in industrial and organizational psychology* (pp. 21-42). Washington, D.C.: American Psychological Association.

Gilchrist, I. (2011). Saccades. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 85-94). Oxford: Oxford University Press.

Gilman, E. R. (2000). *Towards an eye-movement model of music sight-reading* (Volume 1). University of Nottingham, Nottingham, United Kingdom.

Gilman, E. R. (2000). *Towards an eye-movement model of music sight-reading* (Volume 2). University of Nottingham, Nottingham, United Kingdom.

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

Gilman, E., Underwood, G., & Morehen, J. (2002). Recognition of visually presented musical intervals. *Psychology of Music, 30*, 48-57.

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

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

Gordon, E. (1965). The musical aptitude profile: A new and unique musical aptitude test battery. *Bulletin of the Council for Research in Music Education, 6*, 12-16.

Gordon, E. (1970). *Iowa Tests of Music Literacy manual*. Iowa City, Iowa: The Bureau of Educational Research and Service.

Gordon, E. (1971). *The psychology of music teaching.* Englewood Cliffs, N. J.: Prentice-

Hall, Inc.

Gordon, E. (1986). A factor analysis of the Music Aptitude Profile, the Primary Measures of Music Audiation, and the Intermediate Measures of Music Audiation. *Bulletin of the Council for Research in Music Education, 87*, 17-25.

Gregory, R. (2007). *Psychological testing: History, principles, and applications* (5th ed). Boston: Pearson/Allyn and Bacon.

Grutzmacher, P. A. (1985). *The effect of tonal pattern training on the aural perception, reading recognition and melodic sight reading achievement of first year instrumental music students* (Order No. 8514172). Available from ProQuest Dissertations & Theses Global. (303354452). Retrieved from http://search.proquest.com/docview/303354452?accountid=14701

Grutzmacher, P. A. (1987). The effect of tonal pattern training on the aural perception, reading recognition, and melodic sight-reading achievement of first-year instrumental music students. *Journal of Research in Music Education, 35*(5), 171-181.

Gudmundsdottir, H. R. (2003). Music reading errors of young piano students (Order No. NQ88481). Available from ProQuest Dissertations & Theses Global. (305264588). Retrieved from http://search.proquest.com/docview/305264588?accountid=14701

Gudmundsdottir, H. R. (2010a). Pitch error analysis of young piano students' music reading performances. *International Journal of Music Education, 28*(1), 61-70.

Gudmundsdottir, H. R. (2010b). Advances in music reading research. *Music Education Research, 12*(4), 331-338.

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

Gutsch, K. U. (1964). *Objective measurement in instrumental music performance*. Retrieved from http:search.proquest.com/docview/64457442?accountid=14701

Gutsch, K. U. (1966). Instrumental music performance: One approach toward evaluation. *The Journal of Educational Research, 59*(8), 377-380.

Hain, T. C. (2012, May 31). *Eye movement recording devices*. Retrieved March 29, 2016, from http://www.dizziness-and-balance.com

Hall, P. & Macardle, F. (1996). *Piano time sight reading*. Oxford: Oxford University Press.

Halpern, A. R. & Bower, G. H. (1979). *Interference tasks and music reading*. Unpublished manuscript, Stanford University.

Halpern, A., & Bower, G. (1982). Musical expertise and melodic structure in memory for musical notation. *American Journal of Psychology, 95*, 31–50.

Havlicek, L. L. (1968). Effect of delayed auditory feedback on musical performance. *Journal of Research in Music Education, 16*(4), 308-318.

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

Henry, M. (2001). The development of a vocal sight-reading inventory. *Bulletin of the Council for Research in Music Education, 150*, 21-35.

Henry, M. L. (2004). The use of targeted pitch skills for sight-singing instruction in the choral rehearsal. *Journal of Research in Music Education, 52*(3), 206-217.

Henry, M. (2015). Vocal sight-reading assessment: Technological advances, student perceptions, and instructional implications. *Update: Applications of Research* *in Music Education, 33*(2), 58-64.

Hodges, D. (1992). Acquisition of music reading skills. In R. Colwell (Ed.), *Handbook of research on music teaching and learning: A project of the music educators national conference* (pp. 466-471). New York: Schirmer Books.

Hodges, D. A. (2015). The child's musical brain. In G. E. McPherson (Ed.), *The child as musician: A handbook of musical development, 2nd ed.* (pp. 52-66). U. K.: Oxford University Press.

Hooper, S. (1966). The development of a test battery to measure music reading readiness of children in grades three and four. *Journal of Research in Music Education, 14*(2), 115- 125.

Horsley, M., Eliot, M., Knight, B. A., & Reilly, R. (Eds.). (2014). Introduction. *Current trends in eye tracking research* (pp. v-vi). Switzerland: Springer International Publishing.

Hubley, A. M. & Zumbo, B. D. (2013). Psychometric characteristics of assessment procedures: An overview. In K. F. Geisinger, B. A. Bracken, J. F. Carlson, J-I. C. Hansen, N. R. Kuncel, S. P. Reise, M. C. Rodriguez (Eds.), *The APA handbook of testing and assessment in psychology, Vol. 1: Test theory and testing and assessment in industrial and organizational psychology* (pp. 3-19). Washington, D.C.: American Psychological Association.

Hyönä, J. (2011). Foveal and parafoveal processing during reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 819-838). Oxford: Oxford University Press.

Jiang, J., Zhou, Z., Yin, Erwei, Yu, Y., & Hu, D. (2014). Hybrid brain-computer interface (BCI) based on the EEG and EOG signals. *Bio-Medical Materials and Engineering, 24*(6), 2919-2925.

Johnson, T. A. & Evans, C. (2001). *Right@sight: A progressive sight-reading course*. London: Edition Peters.

Kevles, B. (1997). *Naked to the bone: Medical imaging in the twentieth century*. New Brunswick, N. J.: Rutgers University Press.

Killian, J. N. & Henry, M. L. (2005). A comparison of successful and unsuccessful strategies in individual sight-singing preparation and performance. *Journal of Research in Music Education, 53*(1), 51-65.

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

Kopiez, R., & Galley, N. (2002). The musicians' glance: A pilot study comparing eye movement parameters in musicians and non-musicians. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, & J. Renwick (Eds*.), Proceedings of the 7th International Conference on Music Perception and Cognition* (pp. 683-686).

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

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

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

Kornicke, L. E. (1992). *An exploratory study of individual difference variables in piano sight-reading achievement* (Order No. 9301458). Available from ProQuest Dissertations & Theses Global. (304007507). Retrieved from http://search.proquest.com/docview/304007507?accountid=14701

Kostka, M. (2000). The effects of error-detection practice on keyboard sight reading achievement of undergraduate music majors. *Journal of Research in Music Education, 48*(2), 114-122.

Kreader, B., Wester, A., Bell, F., Keveren, P., Breth, N. O. (1996). *Hal Leonard student piano library*. Milwaukee, IL: Hal Leonard Corporation.

Krings, T., Töpper, R., Foltys, H., Erberich, S., Sparing, R., Willmes, K., & Thron, A. (2000). Cortical activation patterns during complex motor tasks in piano players and control subjects: A functional magnetic resonance imaging study. *Neuroscience Letters, 278*, 189–193.

Land, M. F. (1993). Eye-head coordination during driving. Proceedings of the International Conference on Systems, Man and Cybernetics (SMC), France, *3*, 490-494. doi: 10.1109/ICSMC.1993.385060

Latham, A. (2002). *The Oxford companion to music*. Oxford: Oxford University Press.

Lamantia, A-S. (2012). Modification of neural circuits as a result of experience. In D. Purves, G. J. Augustine, D. Fitzpatrick, W. C. Hall, A-S. Lamantia, L. E. White (Eds.), *Neuroscience* (5th ed.). (pp. 537-557). Sunderland, Massachusetts: Sinauer Associates, Inc. Publishers.

Law, L. N. C. & Zentner, M. (2012). Assessing musical abilities objectively: Construction and validation of the profile of music perception skills. *PloS one, 7*(12), 1-15.

L C Technologies. (2015). *Eye tracking for research and analysis*. Retrieved March 29, 2016, from http://www.eyegaze.com

Lee, J. I. (2004). *Component skills involved in sight reading music*. Frankfurt am Main: Peter Lang.

Legore, S. H. (1981). *The development and validation of a test to measure music reading*

Readiness (Order No. 8111872). Available from ProQuest Dissertations & Theses Global. (303102208). Retrieved from http://search.proquest.com/docview/303102208?accountid=14701

Lehman, P. R. (1969). A selected bibliography of works on music testing. *Journal of Research in Music Education, 17*(4), 427-442.

Lehmann, A. C., & Ericsson, K. A. (1993). Sight-reading ability of expert pianists in the context of piano accompanying. *Psychomusicology, 12*(2), 182–195.

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

Lehmann, A. C., & McArthur, V. (2002). Sight-reading. In R. Parncutt & G. E. McPherson (Eds.), *The Science and Psychology of Musical Performance: Creative Strategies for Music Teaching and Learning* (pp. 135–150). Oxford: Oxford University Press.

Lemay, C. (2008). *The measurement, quantification and evaluation of piano students' music reading performance: Comparing methods of assessment* (Unpublished M.A. thesis).

Levy, K. (2001). *Music readers and notation: Investigation of an interactive model of rhythm reading* (UMI No. 3034120).

Linton, O. W. (n. d.) *Medical applications of X rays*. Retrieved March 29, 2016, from University of Stanford Web site: http://www.slac.stanford.edu/pubs/beamline/25/2/25-2-linton.pdf

Lopes da Silva, F. H. (2010). Electrophysiological basis of MEG signals. In P. Hansen, M. Kringelbach, & R. Salmelin (Eds.), *MEG: An introduction to methods* (pp. 1-35). New York: Oxford University Press.

Lowder, J. (1973). Evaluation of a sight-reading test administered to freshman piano classes. *Journal of Research In Music Education, 21*(1), 68-73.

MacKnight, C. B. (1975). Music reading ability of beginning wind instrumentalists after melodic instruction. *Journal of Research in Music Education, 23*(1), 23-34.

MacLeod, C. M. (2005). The Stroop task in cognitive research. From A. Wenzel & D. C. Rubin (Eds.), *Cognitive methods and their application to clinical research* (pp. 17-40). Washington, D. C.: American Psychological Association.

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

Marmor, M. F., Brigell, M. G., McCulloch, D. L., Westall, C. A., & Bach, M. (2011). ISCEV standard for clinical electro-oculography (2010 update). *Documenta Ophthalmologica, 122*, 1-7.

Matthews, B. R. (2008). The musical brain. In G. Goldenberg, & B. L. Miller (Eds.), *Handbook of clinical neurology* (3rd series. Vol. 88. pp. 459-469). Retrieved from http://www.sciencedirect.com/science/handbooks/00729752/88

McConkie, G. W. & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception and Psychophysics, 17*, 578-586.

McDowell, I. (2006). *Measuring health: A guide to rating scales and questionnaires*. (3rd Ed.). New York, NY: Oxford University Press.

McPherson, G. E. (1994). Factors and abilities influencing sight-reading skill in music. *Journal of Research in Music Education, 42*(3), 217-231.

McPherson, G. E. (1995). The assessment of musical performance: Development and validation of five new measures. *Psychology of Music, 23*(2), 142–161.

McPherson, G. E., & Gabrielsson, A. (2002). From sound to sign. In R. Parncutt & G. E. McPherson (Eds.), *The Science and Psychology of Musical Performance: Creative Strategies for Music Teaching and Learning* (pp. 99–115). Oxford: Oxford University Press.

McPherson, G. E., Bailey, M., & Sinclair, K. (1997). Path analysis of a model to describe the relationship among five types of musical performance. *Journal of Research in Music Education, 45*(1), 103–129.

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.

Micheletti, L. (1980). *An assessment of the vertical method and the ensemble approach for teaching sight reading to secondary class piano students* (Doctoral dissertation, University of Miami). (UMI No. 8100741)

Miller, R. E. (1988). *Contributions of selected music skills to music sight-reading achievement and rehearsed reading achievement* (Doctoral dissertation, University of Illinois at Urbana-Champaign, 1988). (UMI No. 8908779)

Mishra, J. (2013). Improving sightreading accuracy: A meta-analysis. *Psychology of Music, 42*(2), 131-156.

Mishra, J. (2014). Factors related to sight-reading accuracy: A meta-analysis. *Journal of Research in Music Education, 6*(4), 452-465.

Mosher, R. M. (1925). A study of group method of measurement of sight-singing. *Contributions to Education*, 194.

Musumeci, O. (1997). The Bentley "measures of musical abilities" as predictors of conservatory students' musical achievement. *Bulletin of the Council for Research in Music Education, 133*, 95-99. Retrieved from http://search.proquest.com/docview/62381140?accountid=14701

Myers, I. B., & Briggs, K. C. (1976). *Myers-Briggs Type Indicator Form F.* Palo Alto, CA: Consulting Psychologists Press, Inc.

Nakada, T., Fujii, Y., Suzuki, K., & Kwee, I. L. (1998). "Musical brain" revealed by high-field (3 Tesla) functional MRI. *NeuroReport, 9*, 3853–3856.

Nolte, J. (2010). *Essentials of the human brain, 1st ed.* Philadelphia: Mosby/Elsevier.

Nunn, J. & Hill, S. (2016). Animated storyboard: What are systematic reviews? Retrieved November 19, 2016, from http://cccrg.cochrane.org/animated-storyboard-what-are-systematic-reviews

Pace, R. (1979). *Music for piano: Robert Pace keyboard approach*. Katonah, NY: Lee Roberts Music Publications.

Pagan, K. A. (1970). An experiment in the measurement of certain aspects of score reading ability (Doctoral dissertation, Indiana University, 1970). *Dissertation Abstracts International*, A 31 (01), 1313.

Palmer, M. (1976). Relative effectiveness of two approaches to rhythm reading for fourth-grade students. *Journal of Research in Music Education, 24*(3), 110-118.

Palmer, W. A., Manus, M., & Vick Lethco, A. (1995). *Alfred's basic piano library*. Sherman Oaks, California: Alfred Publishing Co.

Parker, E. J. (2006). *Piano pedagogy: A practical approach*. Surrey, BC: Longbow Publishing Ltd.

Parncutt, R. & McPherson, G. E. (Eds.). (2002). *The Science and Psychology of Musical Performance: Creative Strategies for Music Teaching and Learning*. Oxford: Oxford University Press.

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.

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.), *12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music* (pp. 792-798). Retrieved from http://icmpc-escom2012.web.auth.gr/sites/default/files/papers/792_Proc.pdf.

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

Peretz, I. (1990). Processing of local and global musical information by unilateral brain-damaged patients. *Brain, 113,* 1185-1205.

Phelps, M. E. (2016). *Positron emission tomography provides molecular imaging of biological processes*. Retrieved June 12, 2016, from http://222.pnas.org/content/97/16/9226.figures-only

Pick, A. D., Unze, M. G., Metz, S., & Richardson, R. M. (1982). Learning to read music: Children's use of structure in pitch notation. *Psychomusicology, 2*(2), 3346.

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

Quadadapt. (n. d.) *Instrument specifications*. Retrieved June 11, 2016, from http://www.quadadapt.com/ins_specs.html

Quinte Health Care. (2016). *Magnetic resonance imaging (MRI)*. Retrieved June 12, 2016, from http://www.qhc.on.ca/mri-p760.php

Randel, D. M. (2002). *The new Harvard dictionary of music*. Cambridge, Mass.: Belknap Press of Harvard University.

Rashbass, C. (1960). New method for recording eye movements. *Journal of the Optical Society of America, 50*(7), 642-644.

Regnault, P., Bigand, E., & Besson, M. (2001). Different brain mechanisms mediate sensitivity to sensory consonance and harmonic context: Evidence from auditory event related brain potentials. *Journal of Cognitive Neuroscience, 13*, 241–255.

Reulen, J. P. H., Marcus, J. T., Koops, D., de Vries, F. R., Tiesinga, G., Boshuizen, K., & Bos, J. E. (1988). Precise recoding of eye movement: the IRIS technique part 1. *Medical and Biological Engineering and Computing, 26*(1), 20-26.

Richards, M. H. (1971). *Threshold to Music*. Belmont, CA: Fearon Publishers.

Roach, D. W. (1982). The development and validation of a test to measure music reading (REVIEW). *Bulletin of the Council for Research in Music Education, 72*, 87-91.

Rowntree, J. P. (1970). The Bentley "Measures of Musical Abilities" – A critical evaluation. *Bulletin of the Council for Research in Music Education, 22*(1), 25-32.

Royal Conservatory of Music (2015). *Piano syllabus*. Mississauga, Ont.: Frederick Harris Music.

Saint Luke's Health System. (2016). *Electroencephalogram (EEG)*. Retrieved June 12, 2016, from http://www.saintlukeshealthsystem.org/health-library/electroencephalogram-eeg

Salis, D. L. (1977). *The Identification and Assessment of Cognitive Variables Associated with Reading of Advanced Music at the Piano* (Doctoral dissertation, University of Pittsburgh, 1977). Retrieved from ProQuest Dissertations and Theses. (UMI No. 789605)

Salis, D. L. (1980). Laterality effects with visual perception of musical chords and dot patterns. *Perception and Psychophysics, 28*(4), 284-292.

Salisbury, F. S. & Smith, H. B. (1929). Prognosis of sight singing ability of normal school students. *Journal of Applied Psychology, XIII*, 425-439.

Salzberg, R. S., & Wang, C. C. (1989). A comparison of prompts to aid rhythmic sight reading of string students. *Psychology of Music, 17*(2), 123-131.

Schlaug, G. (2001). The Brain of Musicians: A Model for Functional and Structural Adaptation. In R.J. Zatorre and I. Peretz (Eds.), *The Biological Foundations of Music* (Annals of the New York Academy of Sciences, Vol. 930, pp. 281–299). New York: The New York Academy of Sciences.

Schleuter, S. L. (1974). Use of standardized tests of musical aptitude with university freshmen music majors. *Journal of Research in Music Education, 22*(4), 258-269.

Schön, D., & Besson, M. (2002). Processing pitch and duration in music reading: A RT-ERP study. *Neuropsychologia, 40*(7), 868-878.

Schön, D., & Besson, M. (2005). Visually induced auditory expectancy in music reading: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience, 17*(4), 694-705.

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

Schön, D., Semenza, C., & Denes, G. (2001). Naming of musical notes: A selective deficit in one musical clef. *Cortex, 37*, 407–421.

Schurmann, M., Raij, T., Fujiki, N., & Hari, R. (2002). Mind's ear in a musician: Where and when in the brain. *Neuroimage, 16*(2), 434–440.

Scripp, L. R. (1995). The development of skill in reading music. *Dissertation Abstracts International Section A: Humanities and Social Sciences*, 56 (6A), 2162.

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

Shaffer, L. H. (1980). Analysing piano performance: A study of concert pianists. In G. E. Stelmach, & J. Requin (Eds.), *Tutorials in motor behaviour*. Amsterdam: North Holland.

Shaffer, L. H. (1981). Performance of Chopin, Bach, and Bartók: Studies in motor programming. *Cognitive Psychology, 13*, 326-376.

Shehan, P. K. (1987). Effects of rote versus note presentations on rhythm learning and retention. *Journal of Research in Music Education, 35*(2), 117-126.

Shuter-Dyson, R., & Gabriel, C. (1981). *The Psychology of Musical Ability* (2nd ed.). Great Britain: Cambridge University Press.

Sidnell, R. (1973). An investigation of the adequacy of the content and difficulty levels of the Iowa Tests of Music Literacy. *Bulletin of the Council for Research in Music Education, 33*, 54-60.

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

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

Sloboda, J. A. (1976b). The effect of item position on the likelihood of identification by inference in prose reading and music reading. *Canadian Journal of Psychology, 30*, 228-36.

Sloboda, J. A. (1977). Phrase units as determinants of visual processing in music reading. *British Journal of Psychology, 68*, 117-124.

Sloboda, J. A. (1978a). The psychology of music reading. *Psychology of Music, 6*, 3-20.

Sloboda, J. A. (1978b). Perception of contour in music reading. *Perception, 7*, 323-331.

Sloboda, J. A. (1983). The communication of musical metre in piano performance. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 35*(2), 377-396.

Sloboda, J. A. (1984). Experimental studies of musical reading: A review. *Music Perception, 2*, 222–236.

Stanzione, M., Grossi, D., & Roberto, L. (1990). Note-by-note music reading: A musician with letter-by-letter reading. *Music Perception, 7*(3), 273-284.

Stelzer, T. G. (1938). Construction, interpretation, and use of a sight reading scale in organ music with an analysis of organ playing into fundamental abilities. *Journal of Experimental Education, 7*(1), 35-43.

Stewart, L., Henson, R., Kampe, K., Walsh, V., Turner, R., & Frith, U. (2003a). Brain changes after learning to read and play music. *Neuroimage, 20*, 71-83.

Stewart, L., Henson, K., Kampe, K. Walsh, V., Turner, R., & Frith, U. (2003b). Becoming a pianist. An fMRI study of musical literacy acquisition. *Annals of the New York Academy of Sciences, 999*, 204–208.

Stewart, L., Walsh, V., & Frith, U. (2004). Reading music modifies spatial mapping in pianists. *Perception & Psychophysics, 66*(2), 183-195.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643-662.

Tatler, B. W. (2014). Eye movements from laboratory to life. In M. Horsley, M. Eliot, B. A. Knight, & R. Reilly (Eds.), *Current trends in eye tracking research*. (pp. 17-35). Switzerland: Springer International Publishing.

Tatler, B. W., Kirtley, C., Macdonald, R. G., Mitchell, K. M. A., & Savage, S. W. (2014). The active eye: Perspectives on eye movement research. In M. Horsley, M. Eliot, B. A. Knight, & R. Reilly (Eds.), *Current trends in eye tracking research*. (pp. 3-16). Switzerland: Springer International Publishing.

Thompson, W. B. (1985). Sources of Individual Differences in Music Sight-Reading Skill (Doctoral dissertation, University of Missouri – Columbia, 1985). *Dissertation Abstracts International, B* 47 (02), 828.

Thostenson, M. S. (1967). The study and evaluation of certain problems in eartraining related to achievement in sightsinging and music dictation. *Bulletin of the Council for Research in Music Education, 11*, 14-35.

Tommis, Y., & Fazey, D. (1999). The acquisition of the pitch element of music literacy skills by 3-4-year-old pre-school children: A comparison of two methods. *Psychology of Music, 27*(2), 230-244.

Treichler, F. R. (2013). Testing musical ability: An American dissenter and some related historical comparisons. *History of the Human Sciences, 26*(5), 48-68.

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.

University of Leicester. (n. d.) *Research facilities*. Retrieved June 12, 2016, from http://www2.le.ac.uk/departments/psychology/research/facilities

Uszler, M., Gordon, S., & Smith, S. M. (2000). *The well-tempered keyboard teacher*

(2nd Ed.). Belmont, CA: Schirmer Books.

Van Nuys, K. & Weaver, H. E. (1943). Memory span and visual pauses in reading rhythms and melodies. *Psychological Monographs, 55*, 33-50.

Ward Technical Consulting. (2015). *Technical: Dual-Purkinje-Image eyetracker*. Retrieved March 29, 2016, from http://www.w63309.wix.com

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

Waters, A. J., Townsend, E., & Underwood, G. (1998). Expertise in musical sightreading: A study of pianists. *British Journal of Psychology, 89*(1), 123–149.

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

Waters, A. J., & Underwood, G. (1999). Processing pitch and temporal structures in music reading: Independent of interactive processing mechanisms? *The European Journal of Cognitive Psychology, 11*(4), 531-553.

Watkins, J. G. (1942). *Objective measurement of instrumental performance*. New York, NY: AMS Press.

Watkins, J. G. & Farnum, S. E. (1954). *The Watkins-Farnum performance scale, form A: A standardized achievement test for all band instruments*. Winona, MN: Hal Leonard Music.

Watkins, J. G. & Farnum, S. E. (1962). *The Watkins-Farnum performance scale, form B: A standardized achievement test for all band instruments*. Winona, MM: Hal Leonard Music.

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

White, L. E. & Hall, W. C. (2012). Lower motor neuron circuits and motor control. In D. Purves, G. J. Augustine, D. Fitzpatrick, W. C. Hall, A-S. Lamantia, L. E. White (Eds.), *Neuroscience* (5th ed.). (pp. 353-374). Sunderland, Massachusetts: Sinauer Associates, Inc. Publishers.

White, L. E. & Hall, W. C. (2012). Modulation of movement by the cerebellum. In D. Purves, G. J. Augustine, D. Fitzpatrick, W. C. Hall, A-S. Lamantia, L. E. White (Eds.), *Neuroscience* (5th ed.). (pp. 417-433). Sunderland, Massachusetts: Sinauer Associates, Inc. Publishers.

White, L. E. & Hall, W. C. (2012). Upper motor neuron control of the brainstem and spinal cord. In D. Purves, G. J. Augustine, D. Fitzpatrick, W. C. Hall, A-S. Lamantia, L. E. White (Eds.), *Neuroscience* (5th ed.). (pp. 375-398). Sunderland, Massachusetts: Sinauer Associates, Inc. Publishers.

Wilkinson, H. P. (1979). Eye movements measured with a solid state photodiode matrix. *Behavior Research Methods & Instrumentation, 11*(4), 429-432.

Wolf, T. (1976). A cognitive model of musical sight-reading. *Journal of Psycholinguistical Research 5*(2), 143-172.

Wong, Y. K. (2010). *Investigations into the role of early visual cortex in expertise reading musical notation*. Retrieved from ProQuest Dissertations and Theses. (UMI No. 3442190)

Wong, Y. K., & Gauthier, I. (2010). A multimodal neural network recruited by expertise with musical notation. *Journal of Cognitive Neuroscience, 22*(4), 695-713.

Wristen, B. (2005). Cognition and motor execution in piano sight-reading: A review of literature. *Update: Applications of Research in Music Education, 24*(1), 44-57.

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

Young, W. T. (1976). A longitudinal comparison of four music achievement and music aptitude tests. *Journal of Research in Music Education, 24*(3), 97-109.

Zhukov, K. (2005). ''Good Sight-readers: Born or bred?'' Paper presented at the 7th Australasian Piano Pedagogy Conference, 1216 July 2005, Adelaide, Australia.

Zhukov, K. (2014a). Evaluating new approaches to teaching of sight-reading skills to advanced pianists. *Music Education Research, 16*(1), 70-87.

Zhukov, K. (2014b). Exploring advanced piano students' approaches to sight-reading. *International Journal of Music Education: Practice, 32*(4), 487–498.

Appendix A: Classified Studies by Category

Test Measurements

Banton, L. J. (1995). The role of visual and auditory feedback during the sight-reading of Music. *Psychology of Music, 23*(1), 3-16.

Bean, K. L. (1938). An experimental approach to the reading of music. *Psychological Monographs, 50*(6), i-80.

Bebeau, M. J. (1982). Effects of traditional and simplified methods of rhythm-reading instruction. *Journal of Research in Music Education, 30*, 107-119.

Betts, S. L., & Cassidy, J. W. (2000). Development of harmonization and sight-reading skills among university class piano students. *Journal of Research in Music Education, 48*(2), 151-162.

Boyle, J. D. (1968). *The effect of prescribed rhythmical movements on the ability to sight read music* (Order No. 6817359). Available from ProQuest Dissertations & Theses Global. (302315007). Retrieved from http://search.proquest.com/docview/302315007?accountid=14701

Demorest, S. M. (1998). Improving sight-singing performance in the choral ensemble: The effect of individual testing. *Journal of Research in Music Education, 46*(2), 182-193.

Drake, C., & Palmer, C. (2000). Skill acquisition in music performance: Relations between planning and temporal control. *Cognition, 74*(1), 1-32.

Eastlund Gromko, J. (2004). Predictors of music sight-reading ability in high school wind players. *Journal of Research in Music Education, 52*(1), 6-15.

Eaton, J. L. (1978). A correlation study of keyboard sight-reading facility with previous training, note-reading, psychomotor, and memorization skills (Doctoral dissertation, Indiana University*). Dissertation Abstracts International*, A 39 (07), 4109.

Elliott, C. A. (1982a). The relationships among instrumental sight-reading ability and seven selected predictor variables. *Journal of Research in Music Education, 30*(1), 5-14.

Elliott, C. A. (1982b). The identification and classification of instrumental performance sight reading errors. *Journal of Band Research, 18*(1), 36-42.

Erlings, B. (1977). A pilot investigation of relationships between elementary keyboard sightreading achievement by music majors in college and selected musical profile tests. *Bulletin of the Council for Research in Music Education*, 14-17.

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.

Grutzmacher, P. A. (1985). *The effect of tonal pattern training on the aural perception, reading recognition and melodic sight reading achievement of first year instrumental music students* (Order No. 8514172). Available from ProQuest Dissertations & Theses Global. (303354452). Retrieved from http://search.proquest.com/docview/303354452?accountid=14701

Grutzmacher, P. A. (1987). The effect of tonal pattern training on the aural perception, reading recognition, and melodic sight-reading achievement of first-year instrumental music students. *Journal of Research in Music Education, 35*(5), 171-181.

Gudmundsdottir, H. R. (2003). Music reading errors of young piano students (Order No. NQ88481). Available from ProQuest Dissertations & Theses Global. (305264588). Retrieved from http://search.proquest.com/docview/305264588?accountid=14701

Gudmundsdottir, H. R. (2010a). Pitch error analysis of young piano students' music reading performances. *International Journal of Music Education, 28*(1), 61-70.

Gutsch, K. U. (1966). Instrumental music performance: One approach toward evaluation. *The Journal of Educational Research, 59*(8), 377-380.

Halpern, A., & Bower, G. (1982). Musical expertise and melodic structure in memory for musical notation. *American Journal of Psychology, 95*, 31–50.

Havlicek, L. L. (1968). Effect of delayed auditory feedback on musical performance. *Journal of Research in Music Education, 16*(4), 308-318.

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

Henry, M. (2001). The development of a vocal sight-reading inventory. *Bulletin of the Council for Research in Music Education, 150*, 21-35.

Henry, M. L. (2004). The use of targeted pitch skills for sight-singing instruction in the choral rehearsal. *Journal of Research in Music Education, 52*(3), 206-217.

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

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

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

Kornicke, L. E. (1992). *An exploratory study of individual difference variables in piano sight-reading achievement* (Order No. 9301458). Available from ProQuest Dissertations & Theses Global. (304007507). Retrieved from http://search.proquest.com/docview/304007507?accountid=14701

Kostka, M. (2000). The effects of error-detection practice on keyboard sight reading achievement of undergraduate music majors. *Journal of Research in Music Education, 48*(2), 114-122.

Lee, J. I. (2004). *Component skills involved in sight reading music*. Frankfurt am Main: Peter Lang.

Lehmann, A. C., & Ericsson, K. A. (1993). Sight-reading ability of expert pianists in the context of piano accompanying. *Psychomusicology, 12*(2), 182–195.

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

Levy, K. (2001). *Music readers and notation: Investigation of an interactive model of rhythm reading* (UMI No. 3034120).

Lowder, J. (1973). Evaluation of a sight-reading test administered to freshman piano classes. *Journal Of Research In Music Education, 21*(1), 68-73.

MacKnight, C. B. (1975). Music reading ability of beginning wind instrumentalists after melodic instruction. *Journal of Research in Music Education, 23*(1), 23-34.

McPherson, G. E. (1994). Factors and abilities influencing sight-reading skill in music. *Journal of Research in Music Education, 42*(3), 217-231.

McPherson, G. E. (1995). The assessment of musical performance: Development and validation of five new measures. *Psychology of Music, 23*(2), 142–161.

McPherson, G. E., Bailey, M., & Sinclair, K. (1997). Path analysis of a model to describe the relationship among five types of musical performance. *Journal of Research in Music Education, 45*(1), 103–129.

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.

Micheletti, L. (1980). *An assessment of the vertical method and the ensemble approach for teaching sight reading to secondary class piano students* (Doctoral dissertation, University of Miami). (UMI No. 8100741)

Miller, R. E. (1988). *Contributions of selected music skills to music sight-reading achievement and rehearsed reading achievement* (Doctoral dissertation, University of Illinois at Urbana-Champaign, 1988). (UMI No. 8908779)

Pagan, K. A. (1970). An experiment in the measurement of certain aspects of score reading ability (Doctoral dissertation, Indiana University, 1970). *Dissertation Abstracts International*, A 31 (01), 1313.

Palmer, M. (1976). Relative effectiveness of two approaches to rhythm reading for fourth-grade students. *Journal of Research in Music Education, 24*(3), 110-118.

Pick, A. D., Unze, M. G., Metz, S., & Richardson, R. M. (1982). Learning to read music: Children's use of structure in pitch notation. *Psychomusicology, 2*(2), 3346.

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

Salis, D. L. (1977). *The Identification and Assessment of Cognitive Variables Associated with Reading of Advanced Music at the Piano* (Doctoral dissertation, University of Pittsburgh, 1977). Retrieved from ProQuest Dissertations and Theses. (UMI No. 789605)

Salis, D. L. (1980). Laterality effects with visual perception of musical chords and dot patterns. *Perception and Psychophysics, 28*(4), 284-292.

Salzberg, R. S., & Wang, C. C. (1989). A comparison of prompts to aid rhythmic sight reading of string students. *Psychology of Music, 17*(2), 123-131.

Scripp, L. R. (1995). The development of skill in reading music. *Dissertation Abstracts*

International Section A: Humanities and Social Sciences, 56 (6A), 2162.

Shehan, P. K. (1987). Effects of rote versus note presentations on rhythm learning and retention. *Journal of Research in Music Education, 35*(2), 117-126.

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

Thompson, W. B. (1985). Sources of Individual Differences in Music Sight-Reading Skill (Doctoral dissertation, University of Missouri – Columbia, 1985). *Dissertation Abstracts International, B* 47 (02), 828.

Thostenson, M. S. (1967). The study and evaluation of certain problems in eartraining related to achievement in sightsinging and music dictation. *Bulletin of the Council for Research in Music Education, 11*, 14-35.

Zhukov, K. (2014a). Evaluating new approaches to teaching of sight-reading skills to advanced pianists. *Music Education Research, 16*(1), 70-87.

Eye-Tracking Measurements

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

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

D'Anselmo, A., Guiliani, F., Marzoli, D., Tommasi, L., & Brancucci, A. (2015). Perceptual and motor laterality effects in pianists during music sight-reading. *Neuropsychologia, 71*, 119-125.

Drai-Zerbib, V., & Baccino, M. T. (2005). L'expertise dans la lecture musicale: Integration intermodale. *L'année psychologique, 205*(3), 387-422.

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

Furneaux, S. & Land, M. F. (1999). The effects of skill on the eye-hand span during musical sight-reading. *Biological Sciences, 266*(1436), 2435-2440.

Gilman, E. R. (2000). *Towards an eye-movement model of music sight-reading* (Volume 1). University of Nottingham, Nottingham, United Kingdom.

Gilman, E. R. (2000). *Towards an eye-movement model of music sight-reading* (Volume 2). University of Nottingham, Nottingham, United Kingdom.

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

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

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

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

Kopiez, R., & Galley, N. (2002). The musicians' glance: A pilot study comparing eye movement parameters in musicians and non-musicians. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, & J. Renwick (Eds*.), Proceedings of the 7th International Conference on Music Perception and Cognition* (pp. 683-686).

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.

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.), *12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music* (pp. 792-798). Retrieved from http://icmpc-escom2012.web.auth.gr/sites/default/files/papers/792_Proc.pdf.

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

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

Sloboda, J. A. (1977). Phrase units as determinants of visual processing in music reading. *British Journal of Psychology, 68*, 117-124.

Sloboda, J. A. (1978b). Perception of contour in music reading. *Perception, 7*, 323-331.

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.

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

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

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

Neurological Measurements

Cappelletti, M., Waley-Cohen, H., Butterworth, B., & Kopelman, M. (2000). A selective loss of the ability to read and to write music. *Neurocase, 6*(4), 321-332.

Fasanaro, A. M., Spitaleri, D. L., Valiani, R., & Grozzi, D. (1990). Dissociation in musical reading: A musician affected by alexia without agraphia. *Music Perception, 7*(3), 259- 272.

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

Nakada, T., Fujii, Y., Suzuki, K., & Kwee, I. L. (1998). "Musical brain" revealed by high-field (3 Tesla) functional MRI. *NeuroReport, 9*, 3853–3856.

Peretz, I. (1990). Processing of local and global musical information by unilateral brain-damaged patients. *Brain, 113,* 1185-1205.

Schön, D., & Besson, M. (2002). Processing pitch and duration in music reading: A RT-ERP study. *Neuropsychologia, 40*(7), 868-878.

Schön, D., & Besson, M. (2005). Visually induced auditory expectancy in music reading: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience, 17*(4), 694-705.

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

Schön, D., Semenza, C., & Denes, G. (2001). Naming of musical notes: A selective deficit in one musical clef. *Cortex, 37*, 407–421.

Schurmann, M., Raij, T., Fujiki, N., & Hari, R. (2002). Mind's ear in a musician: Where and when in the brain. *Neuroimage, 16*(2), 434–440.

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

Stanzione, M., Grossi, D., & Roberto, L. (1990). Note-by-note music reading: A musician with letter-by-letter reading. *Music Perception, 7*(3), 273-284.

Stewart, L., Henson, R., Kampe, K., Walsh, V., Turner, R., & Frith, U. (2003a). Brain changes after learning to read and play music. *Neuroimage, 20*, 71-83.

Stewart, L., Henson, K., Kampe, K. Walsh, V., Turner, R., & Frith, U. (2003b). Becoming a pianist. An fMRI study of musical literacy acquisition. *Annals of the New York Academy of Sciences, 999*, 204–208.

Stewart, L., Walsh, V., & Frith, U. (2004). Reading music modifies spatial mapping in pianists. *Perception & Psychophysics, 66*(2), 183-195.

Wong, Y. K. (2010). *Investigations into the role of early visual cortex in expertise reading musical notation*. Retrieved from ProQuest Dissertations and Theses. (UMI No. 3442190)

Wong, Y. K., & Gauthier, I. (2010). A multimodal neural network recruited by expertise with musical notation. *Journal of Cognitive Neuroscience, 22*(4), 695-713.

Appendix B: Charts of Classified Studies, Alphabetically

A

Title	Eye movement patterns during the processing of musical and linguistic syntactic incongruities
Author	Stephanie Ahken, Gilles Comeau, Sylvie Hébert, Ramesh Balasubramaniam
Year	2012
Purpose	• To investigate the eye movements of readers during visual processing of music and linguistic syntactic incongruities • To study the effect of key signature notation on music-syntactic processing • To examine the role of key signature and accidentals to establish tonality
Measurement	• Eye movements: fixations, regressions in, regressions out, mean fixation duration in target ROI, total trial duration time (that is, reading time) Eye movements were divided into regions of interest (ROIs) For music stimuli, each measure was an ROI
Assessments Used	-
Description	
Subjects	18 experienced pianists
Instruments Used	Piano
Stimuli	Linguistic: 20 short sentences, half of which included syntactic anomaly (either progressive "-ing" affix or past-tense inflection) (Taken from Braze et al., 2002) Musical: Graduate-student composed, 16 short musical phrases in groups of 4. Each phrase had 5, 6, or 7 bars. Half of the sequences were syntactically congruent. Written in 2 forms: key signature OR with accidentals.
Apparatus	*Finale 2007*, Version 2007 .r2 (Make-Music, Inc.) *Toshiba Tecra M4 Notebook* Tablet, with 14.1 inch monitor *Yamaha Disklavier* grand piano *SR EyeLink II* Eye-tracking system (250 Hz sampling rate for both eyes) *EyeLink Data Viewer 1.9.1* (SR Research Ltd.) for data
Method	All subjects completed a short background questionnaire. Subjects sat at piano, and eye tracker was installed and calibrated. There were 2 practice sequences. Subjects were to play each sequence with hands together and without preview, at a tempo of their choice. They also read linguistic stimuli (sentences) aloud at normal speed. Alternated music of 4 random sequences, then text of 10 random sequences and occasional Yes/No comprehension questions, then music, then break for recalibration. Session was 30 minutes.
Results	Found that for incongruent stimuli (both music and text), proportion of fixations was greater, mean fixation duration was longer, and trial duration was longer. First evidence that eye-movement patterns are sensitive to music-syntactic incongruities. Data suggests a shared system for music and language syntax processing. Findings suggest that music with key signature and accidentals are deciphered at equal speed, but unconventional presentation with accidentals has no advantage for congruent over incongruent endings. Incongruent and congruent trials did not differ for accidentals.
Reliability	-
Validity	-
Reference	Ahken, S., Comeau, G., Hebert, S., & Balasubramaniam, R. (2012). Eye movement patterns during the processing of musical and linguistic syntactic incongruities. *Psychomusicology: Music, Mind, & Brain, 22*(1), 18-25.

B

Title	The role of visual and auditory feedback during the sight-reading of music
Author	Louise J. Banton
Year	1995
Purpose	• To examine the role of auditory and visual feedback during the sight-reading of a short piece of piano music • To identify errors that arose due to the removal of feedback • To compare the performances of pianists at different levels of sight-reading practice frequency, general musical experience and sight-reading ability
Measurement	• Role of auditory and visual feedback • Errors in sight-reading • Performances of pianists at various levels of sight-reading practice frequency, general musical experience, and sight-reading ability. Computer notated scores for each performance and these were compared to original scores of test material; examined to establish what errors had occurred during specific performance conditions and for which ability group. Errors categorized into 1 of 10 error types: 1. Accidental errors, 2. Step errors (errors between notes positioned adjacently to each other), 3. Leap errors (interval larger than a third), 4. Omitted note errors, 5. Correction to previous note errors, 6. Incorrect movement assumption errors (i.e. the played music moves in a direction that is directly contrary to written score), 7. Notes played instead of tied errors, 8. False starts (i.e. no errors incurred during first attempt but section is repeated, 9. Repeats due to error/errors during first attempt, 10. Rhythmic errors (i.e. the incorrect on/off timing values of individual notes. Error types also grouped according to following classification: Total errors (types 1-10), Melodic errors (types 1-7), Rhythmic errors (type 10), Specific errors (types 1-9).
Assessments Used	-
Description	
Subjects	15 pianists who had attained Grade 6 or above (Associated Board of Piano Examinations), and who were practicing the piano regularly at the time of the experiment.
Instruments Used	Piano
Stimuli	The 3 pieces were excerpts taken from the Associated Board of Music Sight-Reading Examinations Grade 5 (1985). The material was different for each of the 3 conditions and the pieces were selected on basis of their similar difficulty level and contrapuntal style. The pieces had different time signatures: duple, triple, and quadruple. The material was novel and relatively short (10-12 bars).
Apparatus	*Korg Sampling* Grand electronic keyboard *MIDI* on *Atari SM124* computer *C-Lab* music software for recording and notation of performances 3-ft (92 cms) long screen of wooden board with wrist height allowance of 6 inches (15 cms), placed over keyboard, to prevent subjects from seeing hands *Boss MA-12* external speakers, which were turned off during restricted auditory feedback condition
Method	Each subject sight-read 3 pieces in each of 3 conditions: normal sight-reading, sight-reading when unable to view the keyboard, and sight-reading when unable to hear what was being played. Subjects had 10 minutes to become accustomed to keyboard and to experience performing under each of the respective test conditions. Subjects then received 3 different pieces to sight-read. They were given up to 1 min to view each piece before sight-reading was attempted. They were instructed to adhere to general sight-reading rules of keeping to a steady tempo of their choice, to refrain from repeating sections, and to "keep on going" to the best of their ability. Normal sight-reading condition was attempted first to establish the base-line ability of the subjects, followed by no visual feedback and no auditory feedback conditions.

	Subjects asked to rate their performances as "very bad", "poor", "average", "quite good", "very good" immediately after the piece, and then again after hearing a sound recording of their performances. Subjects also asked to complete a general musical background questionnaire, which served to assign them to their respective testing groups. Based on answers provided on general musical background questionnaire, subjects were assigned to one of three mutually exclusive groups for each of 3 independent scales: 1. General Musical Experience, 2. Frequency of Sight-Reading Practice, 3. Sight-Reading Ability.
Results	Found that visual feedback appears to be a necessary requirement of skilled performance, which facilitates movement accuracy. Found that performances where visual feedback was unavailable resulted in a significantly increased number of adjacent note errors, suggesting that visual feedback is utilized by pianists in order to guide the discrete movements of the hands over the keyboard, but that the degree to which it is relied upon is dependent on the pianist's familiarity with the sight-reading situation itself. With auditory feedback unavailable, sight-reading was indistinguishable from normal sight-reading. Found that general musical experience and sight-reading ability were significant factors in distinguishing between errors which might produce a breakdown in performance and those which will not impede fluency.
Reliability	-
Validity	-
Reference	Banton, L. J. (1995). The role of visual and auditory feedback during the sight-reading of music. *Psychology of Music, 23*, 3-16.

Title	An experimental approach to the reading of music
Author	Kenneth L. Bean
Year	1938
Experiment 1: The Diagnostic Experiment	
Purpose	• To collect data from music readers of all levels of efficiency and with differing background experience.
Measurement	• Recording of every error by experimenter • Span of perception for each type of material Errors caused by looking too high/low were not included
Assessments Used	-
Description	
Subjects	50; 28 of these were men, and 3 were children aged 10, 10, and 11. Subjects ranged from having no knowledge of notation to being professional musicians.
Instruments Used	Piano
Stimuli	Each card showed a staff or stave, with fragments of printed music placed on the cards. Stimuli was divided into 1. Melodic, 2. Polyphonic, 3. Harmonic excerpts. Each category has stimuli increasing in difficulty. Total of 16 sets of stimuli for practice, and another set for testing, and for trial previous to practice. A total of 680 separate cards. Series A, B, C, D, E were only melodic excerpts; Series F, G, H were polyphonic, with some music on one staff and the rest on two staves; Series I, J, K, L were chordal, for one or both hands; Series M was for unclassified material including unusual rhythms, arpeggios, and contemporary music; Series N, O, P were for comparison purposes. Author composed some of the melodic patterns, where other melodies were chosen from work of Bach violin music, John Alden Carpenter, and Samuel Gardner.
Apparatus	Twin tachistoscope, mounted on the piano Cards to present stimuli
Method	During the individual testing sessions, subjects completed preliminary trials to familiarize themselves with the apparatus, and then sight read as many cards from

	each type of series as possible. The subject was told the key signature, time signature, and whether LH, RH, or both hands were to be used. There was a half second between the signal from the experimental and the presentation of the slide. The subject was to play what they saw after one second. If subjects did not have any knowledge of music, they stated the line or space the notes were on, or drew a pattern of notes on a prepared staff template.
Results	Span of Perception: Notes which did not fall near fovea were illegible. For least-skilled group, could ony read one note at a time (know which space or line) Greater training means better reading ability. For skilled readers, some parts of patterns were difficult to see, but skilled readers were more efficient pattern readers
Reliability	-
Validity	-
Experiment 2: The Practice Experiment	
Purpose	• To study the general effects of practice on the ability to organize patterns on the stimuli cards presented
Measurement	• Sight Reading: Each subject received a subjective evaluation following a university grading system, with a rating of A, B, C, D, or E, with plus or minus signs as needed. • Experiment: Persistent errors were noticed and recorded. Type and number of errors recorded.
Assessments Used	-
Description	
Subjects	24 (14 were women, 2 high school girls ages 13 and 14). Almost all technically advanced. All subjects had to have had four years of piano training.
Instruments Used	Piano
Stimuli	Sight Reading: There was a slow tempo piece and a fast tempo piece. They were taken from piano interludes in "Symphonie Espagnole" by Lalo.
Apparatus	Twin tachistoscope, mounted on the piano Cards to present stimuli
Method	Before the experiment, the experimenter had the subjects sight read 2 pieces. They learned about the apparatus and procedure, and given 30 minutes of preliminary trials. The subjects were instructed to play in strict time at a convenient tempo. For the experiment proper, all cards were given in the series except the last few in series F (too difficult) and the entire series of G and H (unsuitable for tachistoscope). These were practiced over the course of weeks. Subjects practiced either 10, 20, or 30 hours. They practiced 3x each week for an hour, except for one week where there were 5 hours of practice completed. Posttest included cards that had not been previously seen and the same two sight reading pieces again.
Results	Practice effects were "general".
Reliability	-
Validity	-
Reference	Bean, K. L. (1938). An experimental approach to the reading of music. *Psychological Monographs, 50*(6), i-80.

Title	Effects of traditional and simplified methods of rhythm-reading instruction
Author	Muriel J. Bebeau
Year	1982
Experiment 1	
Purpose	• To experimentally compare the effects on rhythm-reading accuracy of two methods of rhythm-reading instruction
Measurement	• Criterion test: measured the student's ability to read each of the twelve rhythmic concepts, and measured the student's ability to maintain a steady pulse Scored by three experienced music teachers, independently, for rhythm and tempo

	accuracy. Each judge scored each item at two different times, and compared the scores to resolve discrepancies
Assessments Used	Pretest and posttest: author-developed criterion test for rhythm-reading consisting of 23 items
Description	
Subjects	27 grade three students in private school without previous group instruction in rhythmic pattern-reading, divided into two treatment groups (group 1 = 14 subjects, group 2 = 13 subjects)
Instruments Used	Clapping, and clapping and speaking (Treatment included use of rhythm instruments)
Stimuli	Author-composed rhythm patterns
Apparatus	Tape recorder for performances
Method	15 minutes of group instruction for 18 days. The speech-cue treatment consisted of spoken cues and associated kinesthetic cues The traditional treatment had mathematical cues, and counted pulses aloud as they clapped or played the patterns. The instructional cues at the beginning of each session were different for the groups. Each group learned 12 rhythmic concepts over the weeks of lessons. The students were tested individually and performances were recorded.
Results	The differences between posttest scores were significant and the .01 level, favoring the speech-cue group
Reliability	Inter-rater reliability: .80 to .96
Validity	-
Experiment 2	
Purpose	• To experimentally compare the effects on rhythm-reading accuracy of two methods of rhythm-reading instruction
Measurement	• Criterion test: measured the student's ability to read each of the twelve rhythmic concepts, and measured the student's ability to maintain a steady pulse Judges unknown
Assessments Used	Pretest and posttest: author-developed criterion test for rhythm-reading consisting of 23 items
Description	
Subjects	80 grade three students in a suburban school
Instruments Used	Clapping, and clapping and speaking (Treatment included use of rhythm instruments)
Stimuli	Author-composed rhythm patterns
Apparatus	Unknown
Method	Identical method to experiment 1
Results	30 children were randomly selected from each group for comparison of gain scores Differences between posttest scores were not significant, but differences between gain scores were significant ($p<.05$) in favor of the speech-cue group
Reliability	-
Validity	-
Reference	Bebeau, M. J. (1982). Effects of traditional and simplified methods of rhythm-reading instruction. *Journal of Research in Music Education, 30*, 107-119.

Title	Development of harmonization and sight-reading skills among university class piano students
Author	Steven L. Betts and Jane W. Cassidy
Year	2000
Purpose	• To explore the development of sight-reading and harmonization skills
Measurement	• Sight-Reading skills, including pitch and rhythm errors • Harmonization skills Pre- and post-test videotapes analyzed for pitch and rhythm accuracy. Pitch error defined as any note played incorrectly that was not correct within the beat. Also, any note added or any note deleted from written notation. Rhythm error was any note value played incorrectly, pause of more than 1 beat, addition or deletion of notes, holding through a rest, starting over to and from any point in music. 100 possible points for Harmonization I, 125 for Harmonization II, 75 for Sight-Reading I, and 109 for Sight-Reading II. Final score could have negative number. 33% of total pre-tests analyzed by 2 independent reliability observers
Assessments Used	-
Description	
Subjects	39 non-keyboard music majors enrolled in 6 intact sections of class piano. All students were in second semester of a 4-semester class piano sequence.
Instruments Used	Piano
Stimuli	First harmonization exercise: 8-measure melody in F major and 4/4 time. Subjects added accompaniment, with Roman numeral chord symbols provided above the melody. Second harmonization exercise: 8-measure melody in G major and 4/4 time. There were more hand-position shifts. Primary and secondary triads indicated by Roman numerals. First sight-reading example: 8 measures in G major and 3/4 time. No hand position shifts in RH, and LH used pattern derived from I-IV64-V65 chord progression taught during first semester. Second sight-reading example: 8 measures in A major and 3/4 time. RH had hand-position shifts and finger crossings, including passing the thumb under the fingers. LH had hand-position shifts and used primary and secondary triads in a broken pattern. Accompaniment discs included accompaniment for every exercise in text, with capability to alter tempo and key. They consisted of a rhythm section an/or synthesized "orchestral" parts, with a variety of styles.
Apparatus	Video recorder for pre- and posttest performances *Roland MT-200* sound module for accompaniment discs *MIDI* file discs of accompaniment (Lancaster & Renfrow, 1995)
Method	The 6 classes were taught by four graduate teaching assistants, teaching from an identical syllabus prepared by one of the authors. All classes were 50 minutes long, with 10 minutes of instruction in harmonization and 10 minutes of instruction in sight-reading. Remaining time was spent on transposition, technical exercises, and repertoire. During first week of class, all subjects were videotaped performing pretest of 2 harmonization and 2 sight-reading tasks: First tasks included concepts taught in first semester. Second tasks used material to be learned in second semester (the semester during which study was conducted). For harmonization exercise, subjects added accompaniment of primary chords using one block chord per chord change, and could use root position or inversions. Total of 9 chord changes. For second harmonization exercise, subjects were to play root chord on first beat and the other chord members as a unit of the third beat of each measure. 9 chord changes in total. 3 classes (n=21) randomly assigned to instruction, including classroom practice, using software during harmonization. 3 classes (n=18) were assigned to instruction, including classroom practice, using software during sight-reading. Procedures were identical for pre- and posttest. Posttest occurred during last week of

	semester: Subjects entered room with acoustic piano and videotape machine, and research associate. They stated name and section, and then performed Harmonization I, Harmonization II, Sight-Reading I, and Sight-Reading II. They were allowed to play through harmonization melodies once and allowed 60 s to study each sight-reading exercise. Tempos established on individual basis.
Results	Found that the right hand was more accurate and consistent than the left hand, made less improvement on all tasks, and was slightly more accurate on the harmonization tasks than the sight-reading tasks. Left hand made noticeable gain in accuracy on all tasks and was more accurate on easier tasks than on the more difficult ones
Reliability	r = .97
Validity	-
Reference	Betts, S. L., & Cassidy, J. W. (2000). Development of harmonization and sight-reading skills among university class piano students. *Journal of Research in Music Education, 48*(2), 151-161

Title	The effect of prescribed rhythmical movements on the ability to sight read music
Author	John David Boyle
Year	1968
Purpose	• To determine whether or not empirical observations of music teachers concerning value of movements as aids in teaching rhythm could be validated experimentally 1. What effect, if any will systematic program of rhythm training have on scores on a rhythm sight reading test? 2. What effect, if any, will systematic program of rhythm training have on music sight reading performance ability (measured by WFPS) 3. Will there be a difference between experimental group and control group scores on rhythm sight reading test and WFPS, in regards to rhythm aptitude, intelligence, and respective pretest scores 4. What is relationship between students' scores on rhythm sight reading test and WFPS?
Measurement	• Intelligence • Rhythm aptitude • Rhythm sight reading performance ability: Scored according to criteria for WFPS which pertain to rhythm • Music sight reading performance ability: Scored according to criteria for WFPS All performances recorded and scoring done by panel of 3 qualified music teachers. Scoring done in accordance with specified objective criteria. Scores of 3 judges was averaged. Judges scored, not knowing if it was pre- or post-test.
Assessments Used	Intelligence: The Henmon-Nelson Test of Mental Ability (Rev. Ed., Form A) (Tom & Nelson, 1957) Rhythm Aptitude: rhythm imagery portion of Musical Aptitude Profile (Gordon, 1965) Sight Reading Performance Ability: WFPS (forms A and B) (Watkins & Farnum, 1954) Rhythm Sight Reading Performance Ability: consisted of the rhythm patterns used in the WFPS
Description	
Subjects	24 bands from 22 schools. 24 bands divided into 2 matched groups of 12 bands each. 2 groups were matched for 1. Amount of rehearsal time weekly, years of teaching experience, size of band. Each band assigned to experimental or control group. Randomly selected population of students in each band was tested with pre- and post-testing. 100 subjects from each group tested, or total of 201 subjects (191 at post-test).
Instruments Used	Band Instruments, including percussion (I believe)
Stimuli	Rhythm Sight Reading Test: Rhythm patterns on a single pitch, taken from WFPS.

	No dynamic or expression markings present. Tests on notes F2 and F3 in bass clef; E4, A4, F5 in treble clef. Snare drum notated in usual percussion notation. 2 forms made. Music Sight Reading Test: WFPS Control and Experimental Groups: A Rhythm A Day (Hudadoff, 1963)
Apparatus	-
Method	Each band director completed training with investigator. All tests administered by investigator or research assistant. Performance tests administered during single session. First, rhythm sight reading test. One-half of subjects received Form A of test; other half received Form B. If they received Form A of rhythm test, then given Form B of WFPS. Pretest was the week preceding experimental semester, and continued through first 2 weeks of semester. Posttests administered during final two weeks of semester. Total of 14 weeks in between for training. For each week, each director spent 10 minutes in 3 rehearsals per week with specified materials and techniques I teaching of rhythm reading. Total of 30 minutes/week. Experimental: Used same materials as control group, but also were to: 1. Listen to recordings of music to recognize beat (including Latin American dance music, marches, overtures), 2. Marked times to underlying beat, 3. Tapped foot to mark beat, 4. Clapped rhythms while tapping beat with foot, 5. Played patterns on single note while tapping beat with foot Control: Spent same amount of time with rhythm training. Control students were not allowed to use any bodily movements with the beat.
Results	Both control and experimental groups made statistically significant gains in scores on rhythm sight reading test and WFPS. However, scores by experimental group on each criterion measure were significantly higher than control group. Mean gain by experimental group on rhythm sight reading test was significant at .01 level of confidence. Significant gains in scores for experimental group with WFPS. With rhythm aptitude, intelligence, and the respective pretest scores held constant, the experimental group's scores on both rhythm sight reading test and WFPS were significantly higher (.01 level) than control group. Also found that rhythm-reading training helped subjects to improve their music sight reading scores as measured by WFPS. Also rhythm-reading training with bodily movement as an aid is significantly more effective than methods that do not employ this.
Reliability	The Henmon-Nelson Test of Mental Ability = .91 (grade 8) and .94 (grades 7 and 9) Musical Aptitude Profile = .86 WFPS = .87-.94
Validity	The Henmon-Nelson Test of Mental Ability = most coefficients .50-.70 Musical Aptitude Profile = .60
Reference	Boyle, J. D. (1968). The effect of prescribed rhythmical movements on the ability to sight read music (Order No. 6817359). Available from ProQuest Dissertations & Theses Global. (302315007). Retrieved from http://search.proquest.com/docview/302315007?accountid=14701

Title	Music rehearsal increases the perceptual span for notation
Author	Douglas D. Burman and James R. Booth
Year	2009
Purpose	• To examine the effect of music rehearsal on the effective visual field (Perceptual Span) • To examine the potential relationship between the perceptual span and maximal performance speed
Measurement	• Perceptual effects of music expertise and familiarity through rehearsal 1. Perceptual span measured with tachistoscopic error-detection task 2. Effect of music rehearsal on perceptual span 3. Effects of musical skill on perceptual span
Assessments Used	-
Description	
Subjects	11 skilled (collegiate musical instruction) and 10 less-skilled (n more than 5 years musical instruction) musicians; all adults, ages 20-45
Instruments Used	-
Stimuli	Rehearsal Passages: "Easy" passage was original, 1-page, with 29 measures of single melodic line in D Major. Had 8th notes in 4/4 time. Contained total of 150 notes. "Difficult" passage was unfamiliar (Liszt's Etude No. 5), 1-page, with 15 measures from. Had 32nd notes in 2/4 time. Parts of passage had single melodic line in treble clef, accompanied by rests in bass clef. Contained total of 176 notes. Experimental Stimuli: Single melodic line from rehearsed passage was target sequence, consisting of 18 notes of equal duration. Experimental stimuli were 3-18 notes from this target sequence. A single note was replaced in variant sequences. For each stimuli length, 4 variant sequences were used. 1/3 of trials: experimental stimulus matched target sequence. Remaining trials: single note from target sequence was replaced with variant note.
Apparatus	Tachistoscope: head-mounted infrared eye tracker (*Iscan*, 120 Hz sampling rate) Electronic keyboard Videotape and recorder for additional analysis for 4 subjects *Encore* music notation software for generating stimuli *Macintosh* laptop computer via *PsyScope* for presentation and recording
Method	One of 2 unfamiliar musical experts assigned as challenging rehearsal passage. Before the first rehearsal, subjects read and rehearsed the passage one time as quickly as possible, without stopping. Immediately after the play-through, they were tested on their ability to recognize a sequence of notes from the music. Subjects were then tested after the first, tenth, twentieth rehearsal. Rehearsal sessions occurred 2-3 days apart. Group 1: Rehearsed 2x each session for 10 sessions Group 2: Rehearsed 3-4x each session for 7 sessions Group 3: Rehearsed 4-5x each session for 4 sessions Perceptual Task: Sequence of notes from rehearsed passage flashed onscreen at left edge of a staff. Same musical phrase always tested, with the sequence differing from rehearsed passage by one note (at most). Subjects pressed one key if sequence of notes matched, and another key if it did not. They were to respond as quickly as possible. Key signature of target passage presented once at beginning of each test session, followed by the trials. Each trial had appearance of a musical staff, followed 500 ms later by experimental stimulus. Experimental stimulus presented for 200 ms, then replaced by empty musical staff. Perceptual test consisted of 245-270 trials over span of 12.25-13.50 minutes. Each trial duration was 3 s.
Results	Found perceptual span of 4-5 notes (consistent with previous reports). Rehearsal produced incremental increases in perceptual span for rehearsed music. Found that perceptual span/performance speed correlations suggest that increases in perceptual

	span during early stages of rehearsal may help mediate improvements in music performance, especially for difficult passages. Finally, skilled musicians initially perceived more notes; however, with challenging passages, the difference in perceptual span was modest and disappeared after 20 rehearsals. Findings are similar to combined "EHS plus perceptual span" demonstrated with moving window technique: perceptual span of 4 notes for less skilled, 5 notes for skilled with difficult passage, and 7 notes for skilled with easier passage. Findings suggest that size of perceptual span with skilled musicians depends on familiarity with local note patterns in music.
Reliability	-
Validity	-
Reference	Burman, D. D., & Booth, J. R. (2009). Music rehearsal increases the perceptual span for notation. *Music Perception: An Interdisciplinary Journal, 26*(4), 303-320.

Title	A selective loss of the ability to read and to write music
Author	Marinella Cappelletti, H. Waley-Cohen, B. Butterworth, M. Kopelman
Year	2000
Purpose	This is a case study of a professional musician who has sustained left posterior temporal lobe lesion and a small right occipito-temporal lesion in episode of hemorrhagic encephalitis, which resulted in selective impairment in reading, writing, and understanding musical notation. To further understand the processing of symbolic notations. To answer the question, "To what extent do these systems (specifically, notation-alphabetic, logographic scripts, numerals, musical notation) function independently in the brain?"
Measurement	Measured three areas: 1. Is patient's music reading deficit independent of ability to read letter, words, numerals, other symbols? 2. Is patient's music reading deficit indicative of a more general impairment of musical ability? 3. Is patient's music reading deficit confined to the 'segmental' parameter or affected equally the understanding of the 'suprasegmental' marks?
Assessments Used	(See neuropsychological assessment) Letters and Words – NART Test (Nelson and Willison, 1991), Syllable Reading Test (Kay et al., 1992), Regular and Irregular Word Reading Test (Kay et al., 1992), Non-Word Reading (Kay et al., 1992), The Allport Test (Allport, 1977), Western Aphasia Battery (Kertesz, 1982; sentence reading subtests administered)
Description	
Subjects	Case study of a 51-year-old professional musician. At age 45, suffered episode of acute encephalitis with coma for 2 days. Reports that she has visual problems, episodic memory impairments, naming deficits, problems with mental calculation, problems with reading and written. There had been some improvement, but the problems persisted enough to affect her music playing
Instruments Used	Piano
Stimuli	**Tests of Reading:** 1. Musical Notes – Based on special tasks used in formal music examinations (Specimen Sight-Reading Test, The Associated Board of the Royal Schools of Music; Levels 1-5 used). The single notes and famous scores were written on individual cards. Subject-composed scores presented in their original form 2. Letters and Words – (See Assessments Used for specifics) 3. Numerals – Set of 100 Arabic numbers: 10 single-digit numerals, 10 numerals between 10-19, 20 numerals between 20-99, 30 numerals between 100-999, 30 numerals between 1000-9999. Also 10 decimals and 10 fractions to be read aloud 4. Musical Symbols – Based on special tasks used in formal music examination (Specimen Sight-Reading Test, The Associate Board of the Royal Schools of Music) 5. Other Symbols – Set of non-alphanumeric symbols contained familiar and unfamiliar symbols: 12 zodiac symbols, 6 punctuation marks, 4 playing cards, 4 arithmetical signs, 4 chemical signs **Tests of General Musical Ability:** 1. Visual Perception Tests – 14 pairs of musical notes. Half were physically identical and half were different 2. Auditory Recognition of Melodies – 16 melodies progressively increasing in length. The 3 alternative melodies (to the target) were matched for rhythm or pitch. One corresponded to target melody, one was very close (one note different), one was very different. When stimuli matched for pitch, rhythm changed, and vice versa

	3. Visual Recognition of Melodies – a. Matching for note duration on staff: 5 different target musical notes on staff, with 4 other distractor notes. Target notes were visually present 5 times. The distractor notes were visually similar (2) or were visually dissimilar (2). Total of 25 stimuli, b. Matching for note duration in isolation: No stave, and subject recognized target notes by choosing between 5 notes presented for each target. Total of 25 stimuli, c. Matching for pitch: Subject identified note with same pitch as target in a 3-choice forced-choice recognition format. 30 stimuli used 4. Tests of Transcoding Musical Notes – (See Method for specifics) 5. Tests of Writing Music – Set of 12 musical notes (6 in treble clef; 6 in bass clef) and 10 musical symbols. Same as used in reading tests 6. Tests of Music Recall – (See Method for specifics) 7. Tests of Understanding Musical Notes' Meaning – (See Method for specifics) 8. Tests of Playing Music – (See Method for specifics)
Apparatus	Only what is needed for each subtest; namely, pencil, paper, slides, musical scores
Method	***Neuropsychological Portion*** *1. Neuropsychological assessment: Verbal IQ, performance IQ, Wechsler Adult Intelligence Scale (WAIS-R) (Wechsler, 1981).* *2. Language examination: Graded Naming Test (McKenna and Warrington, 1983), Western Aphasia Battery (Kertesz, 1982)* *3. Number and calculation examination: oral and written: ability to count out aloud, knowledge of number order, transcoding numbers from one numerical format to another (5 to five), comparing pairs of numbers for quantity, ability to define the four arithmetical operations, knowledge of arithmetical signs, multidigit addition and subtraction (The Graded Difficulty Arithmetic Test; Jackson and Warrington, 1986)* *4. Memory examination: revised Wechsler Memory Scale (Wechsler, 1987), Recognition Memory Test (Warrington, 1984), Rey-Osterrieth Complex Figure Test (Rey, 1941; Osterrieth, 1944), Pyramids and Palm Trees Test (Howard and Patterson, 1992)* *5. Attention and concentration examination: 'Test of Everyday Attention' (Roberston et al., 1994)* *6. Executive function examination: Modified Card Sorting Test (Nelson, 1976), F. A. S> Test (Benton and Hamsher, 1989), Trail Marking Test (Reitan, 1971)* **Experimental Investigation** **Tests of Reading:** 1. Musical Notes – Subject read series of single musical notes (treble and bass clefs) and musical scores (half famous and half subject-composed before illness). Each was presented in a random order on cards 2. Letters and words – Presentation of letters of alphabet, syllables, words, non-words, sentences (See Assessments Used for specifics) 3. Numerals – Read aloud (See Stimuli for specifics) 4. Musical Symbols – Patient read series of symbols presented by experimenter or defined by experimenter 5. Other Symbols – Subject read aloud a series of non-alphanumeric symbols **Tests of General Musical Ability:** 1. Visual Perception Tests – Subject discriminated musical notes as "same" or "different". One half of pair was presented, followed immediately by second half. Patient was to ignore the musical value and judge based on physicality only 2. Auditory Recognition of Melodies – Patient listened to a target melody and 3 alternatives. Patient was to judge which of the 3 alternatives corresponded to the target 3. Visual Recognition of Melodies – Subject was to recognize musical notes

	presented with or without the context (namely, the score) (See Stimuli for specifics) 4. Tests of Transcoding Musical Notes - The set of musical notes was presented in three different formats: musical notes, note names from A-G, notes played on a piano by the experimenter. 5 tests: a. Matching spoken note names (A-G) to notes on the stave – subject matched the note names read aloud by experimenter to corresponding note. Each target note presented with 2 alternatives: a musical note very different from target and one close to target, b. Matching written note names (A-G) – Same as a. but with written notes, c. Matching played notes to notes on the stave – subject matched each target note played to corresponding written note. Notes were played first to give reference frame, d. Matching played notes to written note names (A-G) – Same as c. but matched to corresponding written musical note name, e. Matching spoken note names (A-G) to played musical notes – a complement to d. where subject matched each target note to the musical note played on the keyboard of the piano 5. Tests of Writing Music – Subject wrote single musical notes and musical symbols to dictation. Also copied single notes and musical symbols. 4 conditions: a. Writing single musical notes to dictation, b. Writing musical symbols to dictation, c. Copying musical notes, d. Copying musical symbols 6. Tests of Music Recall – a. Memory for single notes: Presented with 24 musical notes individually played by examiner and asked to sing back each note immediately, b. Memory for melodies: Presented with series of 16 short melodies composed of a few musical notes individually played by the examiner. Subject listened and asked to sing it immediately 7. Tests of Understanding Musical Notes' Meaning – Subject judged whether 2 musical notes (one in bass and one in treble) represent the same note or not. In have of the 14 pairs, notes had same value. In other half, had a different value. Subject did not name the notes but only identified them according to the value 8. Tests of Playing Music – Subject played individual notes by score and by the examiner naming them
Results	**Tests of Reading:** 1. Musical Notes – Subject was not able to read any music 2. Letters and Words – Subject able to read letters, words, sentences normally 3. Numerals – Subject able to read Arabic numerals 4. Musical Symbols – Able to read the musical symbols. This suggests reading note deficit is independent of ability to read symbols 5. Other Symbols – Correctly read 29 of 30 symbols **Tests of General Musical Ability:** 1. Visual Perception Tests – Able to judge all pairs correctly 2. Auditory Recognition of Melodies – Able to recognize 13 of 16 melodies. Matched for pitch was better 3. Visual Recognition of Melodies – Subject able to match notes for duration. Did not perform well for pitch matching. Subject could have used some general knowledge of notation without an understanding of the specific duration and pitch 4. Tests of Transcoding Musical Notes – If task involved musical notes, the performance was at floor (0% correct). If the task involved only note names or sounds corresponding to each musical note, then performance was at ceiling (100% correct) 5. Tests of Writing Music – Clear pattern: If task involved writing musical notes to dictation, then 0% correct. If task involved writing musical symbols to dictation or copying both musical notes and symbols, then 100% correct. This suggests that subject's deficit is not attributed to impaired visual or motor processes. Also suggests that musical notes and symbols can be

	thought of as two distinct subsystems that are processed separately 6. Tests of Music Recall – 100% correct. Suggests that musical memory is intact and deficit cannot be attributed to musical memory impairments 7. Tests of Understanding Musical Notes' Meaning – Subject judged all notes as different even when they were the same. Therefore, this was performing at a chance level. This inability was consistent with not understanding the notes' meaning as the subject could also not name them 8. Tests of Playing Music – Subject could not play the notes when written on the score, but could play them when named by the examiner General Summary – Able to perceive and recognize notation, and could match notes and copy them. Could not understand notes as they represented element in the musical system, name them, or name their value. Could not write musical dictation. The fact that the tests were at ceiling or floor, implied total impairment or spared ability. The impairment appears to be specific, isolated and format-dependent deficit. The type of script (i.e. musical note) determines the subject's ability to read
Reliability	-
Validity	-
Reference	Cappelletti, M., Waley-Cohen, H., Butterworth, B., & Kopelman, M. (2000). A selective loss of the ability to read and to write music. *Neurocase, 6*, 321-332.

D

Title	Perceptual and motor laterality effects in pianists during music sight-reading
Author	Anita D'Anselmo, Felice Guiliani, Daniele Marzoli, Luca Tommasi, Alfredo Brancucci
Year	2015
Purpose	• To test the relative ability of the two hemispheres in music sight-reading • To assign to one of the main processes of music reading the cause of a possible asymmetry
Measurement	• To analyze perception and early processing • To test (separately) the perceptual input from the motor output processes by means of comparison between musical and verbal notation conditions, and between left hand, right hand, and 2-hands condition • Errors and response times (RT) (only considered RTs of correct responses) Errors detected by 2 trained pianists in 'blind mode', where they compared performance of each subject to actually presented notation. Error was calculated as percentage of correctly played notes within the ones composing the stimulus.
Assessments Used	-
Description	
Subjects	46 skilled pianists
Instruments Used	Piano
Stimuli	Three types: A single note or chord in upper and lower staff (2-hands), note or chord only in lower staff (LH), note or chord only in upper staff (RH). Stimuli presented in classical musical notation with notes on staves, and a verbal notation where note name and chord type were written in capital letters. Total of 44 stimuli for each condition. Each stimulus was presented in left visual field (LVF) and right visual field (RVF). Musical stimuli were on 5-line staff. Verbal notation was written with note names (Do, Re Mi, etc.) accompanied by MAG (Major) or min (minor) to indicate triads
Apparatus	Tachistoscope *Yamaha PSR-E413* digital keyboard *Microsoft Visual Basic* software for stimuli *Philips SHP5400* Stereo Headphone for listening to performance *Fujitsu Siemens* Computer, *Intel (R) Pentium (R)* processor computer *Cubase* (v. SX 1.01, Steinberg Cubase) for recording performance *GoldWave* (V.5.25, GoldWave Inc.) for recording audio signals *Sony DCR-HC17E* video camera monitored eye-movements
Method	Experiment began with training. Subjects performed 2 sessions – one with musical notation and one with verbal notation. Subjects sat at keyboard and fixated at cross in center of screen. They were not to move eyes during experiment. Each trial presented a tachistoscopic stimulus in LVF or RVF. Presented for 180 ms (verbal) or 300 ms (notational), with inter-stimulus interval of 6000 ms. Subjects were to play the stimulus as quickly as possible on the keyboard. Each session had 88 trials and lasted 10 minutes.
Results	Found the expected "Simon" effect – producing faster responses for stimuli performed with hand ipsilateral (belonging to or occurring on the same side of the body) to side of presentation. Observed tendency in favor of the left hemisphere (right visual field (RVF)) only with single-hand tasks. For two-hand task, observed an opposite asymmetry with musical notation. The bass clef had an unexpected role of penalizing LH performance, particularly with LVF stimuli. There is a variegated pattern of hemispheric asymmetries in sight reading, which depend on stimulus coding and motor output type. RH asymmetry with 2-hands playing, while one hand is more leftward lateralized
Reliability	-

Validity	-
Reference	D'Anselmo, A., Guiliani, F., Marzoli, D., Tommasi, L., & Brancucci, A. (2015). Perceptual and motor laterality effects in pianists during music sight-reading. *Neuropsychologia, 71*, 119-125.

Title	Improving sight-singing performance in the choral ensemble: The effect of individual testing
Author	Steven McGregor Demorest
Year	1998
Purpose	• To examine the effect of individual testing, in conjunction with group instruction, on students' sight-singing skills
Measurement	• Rhythmic errors, pitch errors, tempo changes, repeated notes, beginning again from the start Scored by two evaluators trained to use the scoring system
Assessments Used	Pretest: Musical background questionnaire Pretest and Posttest: Test consisting of one major and one minor melody previously unseen
Description	
Subjects	306 subjects from beginning and advanced choirs of six high schools In each school, two intact choirs were assigned to the experimental or control condition Each condition had three beginning and three advanced groups
Instruments Used	Voice
Stimuli	Pretest and Posttest: Melodies adapted from Music for Singing (Ottman, 1967); each example was 15 notes in length Treatment: Major melodies from Oxford Folk Song Sight=Singing Series (Crowe, Lawton, & Whittaker, 1961) Volumes 1 and 2
Apparatus	Two tape recorders for performances
Method	Pretest and Posttest: Subjects entered practice room and read written instructions to turn on recorder, give name, then listen to other recorder for verbal instructions. They used folder for their voice part and heard the tonic chord and first note of the melody. After 30 seconds of practice, they were given the tonic chord and first pitch again and were instructed to sing example Experimental Treatment: Individual testing in addition to regular group instruction. Treatment began 4 weeks after pretest, administered once a month for 3 times during a semester. Posttest administered at beginning of next semester. During semester, students practiced melodies on their own, and were tested after a week on a melody of portion of a melody from their practice sheets. They recorded their singing for the test. They received comments on successes and problems. Control Groups: Only group instruction but using the same melodies as the experimental group in a group setting only
Results	Experimental group had significant gain in performances of the major melodies Suggests that individual testing can help to improve individual performance Lack of significant treatment effect of the minor melodies
Reliability	Inter-evaluator reliability .97
Validity	-
Reference	Demorest, S. M. (1998). Improving sight-singing performance in the choral ensemble: The effect of individual testing. *Journal of Research in Music Education, 46*(2), 182-193.

Title	Expertise In Musical Sight-Reading: Intermodal integration
Author	Veronique Drai-Zerbib and M. Thierry Baccino
Year	2005
Purpose*	• To demonstrate the status of mental structures for recall of musical scores and reconstruction of musical information, and to show that these structures are amodal; they are constructed regardless of input modality (visual or auditory)
Measurement*	Measurements were done with a method based on reflection of infrared light in the center of the cornea to follow the eye's movements. These sets of coordinates given by the recording of ocular movements allow analysis of several variables: • Fixation positions • Trajectory of sight-reading (progressive and regressive saccades) • Temporal progression of the eye (duration of fixations, duration of progressive and regressive fixations) This analysis was done while including the following variables: the type of musician (expert or not), the absence or presence of slurs on the scores heard or read before playing, and the regions of interest preselected where the ocular fixations occurred.
Assessments Used	Pretest: The level of musical reading of each participant was evaluated by the number of ocular fixations on a musical score. These participants were then divided statistically into two groups.
Description	
Subjects	20 expert pianists (evaluated as high performing musical sight-readers) and 7 non-expert pianists (evaluated as weak musical sight-readers)
Instruments Used	Piano.
Stimuli*	16 musical partitions, each containing 4 lines divided in 2 parallel staffs, one for the melody played by the right hand and one for the bass played by the left hand. Presented in two different versions: one contains slurs, the other has none (in both visual modality and auditory modality). These partitions contained scores from Czerny, Bartók, Scarlatti, or ones composed solely for the purpose of the experiment
Apparatus	Digital Piano keyboard Computers using *Eye-Gaze* software to register and trace ocular movements filmed with infrared light fixed on the eye *Sound Forge*™ software for analyzing recordings done experimentally
Method	Participants listened to a series of musical fragments (either with slurs or none). They then read a sheet with the musical fragments (with or without slurs). The participants played 16 partitions on a digital piano keyboard by sight-reading the provided musical sheets. The ocular measuring apparatus was recalibrated after 4 lines each time.
Results*	There was no significant effect from presence of slurs on the number of total fixations. However, there was an effect related to the level of expertise. Experts had a significantly smaller average number of fixations than non-experts. There was also a significant effect regarding the area of the page that was looked at with level of expertise. Regardless of region, non-experts had a significantly greater average of fixations. There was a significant relation between part of the page looked at and presence of slurs. A significant interaction was found between regions looked, presence of slurs, and expertise. This relationship implies that when readers are non-experts, the average number of fixations is greater from regions dl to d4 (melodic line) when slurs are absent than when present. This difference does not happen with experts. The effect wasn't present for the bass part. Experts made significantly less fixations than non-experts regardless of slurs and fixated, on average, less than non-experts. Non-experts fixated, on average, faster when slurs are present while the speed of fixation remains the same in both conditions for experts. Experts have a significantly smaller average number of regressive fixations than non-experts and show a relationship with types of areas

	looked at. The average number of regressive fixations is less significant when musical material contains slurs than when they are absent. There is a significant interaction between presence of slurs and level of expertise. The average number of regressive fixations is inferior in non-expert when slurs are present while the difference is not significant in experts. The relationship between expertise and the region of the page read, as well as the relation between the presence of slurs and the region read were also significant. The double interaction between the presence of slurs, parts of the page looked at, and level of expertise was also significant. The average number of regressive fixations is more important when the staff contains slurs than when absent only when the reader is non-expert. There were no effects for the presence or absence of slurs and expertise for the number and duration of progressive fixations. Experts could play the scores with slurs even if they weren't written on the sheet, while the non-experts had much difficulty and didn't include slurs when not written on the sheet.
Reliability	-
Validity	Contains results that are not statistically significant
Reference	V. Drai-Zerbib, M. Thierry Baccino. (2005). L'expertise dans la lecture musicale : intégration intermodale. *L'année psychologique, 105*(3), 387-422.
Note	*This chart was written by translation from the original French.* *'*' denotes direct translations.*

Title	Sight-reading expertise: Cross-modality integration investigated using eye tracking
Author	Veronique Drai-Zerbib, Thierry Baccino, and Emmanuel Bigand
Year	2012
Purpose	• To demonstrate the impact of a written code that supplies visuomotor cues, or fingering on eye movements • To validate hypothesized modal independence of expert pianists, meaning that they can retrieve information whether they perceive information visually or auditorily
Measurement	• Eye movement, with eye-tracking data based on 9 areas: the clef area, the four treble clef measures played with RH, four bass clef measures played with LH • Four dependent variables analyzed for each of the above areas: first-pass fixation duration, second-pass fixation, the probability of re-fixation, and playing errors
Assessments Used	-
Description	
Subjects	25 piano students or piano teachers from National Conservatory of Music in Nice, France Divided into 2 expertise groups: 15 experts, with 12+ years of practice, and 10 non-experts, with 6-8 years of study
Instruments Used	Piano
Stimuli	36 piano excerpts, at four measures long each, taken from classical tonal repertoire Each excerpt had 3 versions for each fingering condition: Difficult, easy, or no fingering
Apparatus	*Finale* software for excerpts Steinway piano played by teacher and recorded for the auditory phase *TOBII Technology 1750* eye-tracking system for ocular data, connected to 2 PC computers 17" screen for music display *Plantronics* headphones for excerpt heard by subjects *Yamaha P120* 88-key electronic piano for performance Portable *Maxdata* PC computer for recording performances *E-prime* software for experiment programming
Method	The eye tracker was calibrated for each subject Practice phase with 3 excerpts, and followed by the experiment Experimental phase included: 1. Fixation cross with or without listening, 2. Reading

	of excerpt with easy, difficult, or no fingering, 3. The words “Phase 2” on screen, 4. Sight reading excerpt performed, 5. Experimenter pressed space bar for next excerpt Repeated this for all 36 excerpts, in random order Eye movements and musical production recorded for each excerpt
Results	Found that the upper staff (RH) was processed more than lower staff (LH) except when music was heard before, in which case suggests that the musicians processed the RH during listening Clef area was passed over quickly, but looked back at more when they first heard the music, suggesting validation of key Experts processed the difficult fingerings on the second pass Experts made use of listening phase to determine the suitability of the proposed fingering Non-experts focused on excerpts containing difficult fingering; they were trying to resolve difficult fingering problems while experts would often ignore difficult fingering and rely on their training During the second-pass, experts looked back less at the music Experts used the first-pass to anticipate potential motor problems, which suggests that just viewing the score facilitates planning and preparation for motor execution, where non-experts do not seem to have this cross-modal integration ability Overall, experts show cross-modal integration and efficient processing retrieved from memory
Reliability	-
Validity	-
Reference	Drai-Zerbib, V., Baccino, T., & Bigand, E. (2012). Sight-reading expertise: Cross-modality integration investigated using eye tracking. *Psychology of Music, 40*(2), 216-235.

Title	Skill acquisition in music performance: relations between planning and temporal control
Author	Carolyn Drake and Caroline Palmer
Year	2000
Purpose	• To track cognitive changes that result from long-term and short-term learning in music performance
Measurement	• Errors were identified by computer and coded according to a system similar to those used with speech errors • 5 error types: pitch errors, duration errors, pitch/duration combinations, corrections, and pauses Results were verified by the notated transcriptions of two musician listeners, and the listeners’ coding was adopted when there was a discrepancy
Assessments Used	Pretest: Interview for musical background
Description	
Subjects	60 subjects in 5 groups of 12. 48 were children aged 7-16 and 12 were adults, aged 20-30. The classification was based on the sight-reading level, and age.
Instruments Used	Piano
Stimuli	Author-composed: Three short Western tonal musical pieces, varied in difficulty. Simple meter, melody, and accompaniment. Piece 1: single-voice, 8 measures, 2 phrases. Piece 2: melody and 2 accompanying voices, with chordal accompaniment, 8 measures, 2 phrases. Piece 3: melody and 2 accompanying voices, based on French folk tune, ABA form, each section was 16 measures
Apparatus	Upright *Yamaha Disklavier* acoustic piano for performances Computer for monitoring *Disklavier* errors
Method	Subjects completed interview and performed a piece on piano for comfort They performed the stimuli, up to 5 times each, but moved on to the next piece if they played it 2 times without error When they established the appropriate piece, they played it 5 times

	They performed a short, unrelated task, and then played the final piece one more time, and a computer played back the recording They were given instructions to emphasize a particular part of the piece, which had no effect on the results, and the piece was performed 5 more times
Results	Found that tempo and accuracy improved with skill level and practice in early trials, timing/pitch accuracy trade-off with the focus on time across skill levels and practice, frequent joint pitch/duration errors than expected. Found that ability to maintain temporal continuity improved with skill level and practice, respect of underlying beat and sensitivity to metrical structure improved with skill level. Found proportion of anticipatory errors increased with skill level and practice, range of planning increased with skill level, and increased as temporal disruptions decreased, with increased skill level.
Reliability	-
Validity	-
Reference	Drake, C., & Palmer, C. (2000). Skill acquisition in music performance: Relations between planning and temporal control. *Cognition, 74*(1), 1-32.

E

Title	Predictors of music sight-reading ability in high school wind players
Author	Joyce Eastlund Gromko
Year	2004
Purpose	• To investigate relationships among music sight-reading as measured by the Watkins-Farnum Performance Scale and tonal and rhythmic audiation, visual field articulation, spatial orientation and visualization, and academic achievement in math concepts and reading comprehension
Measurement	• Tonal and rhythmic audiation • Spatial Orientation • Visual Field Articulation • Spatial Visualization • Sight-Reading Performance
Assessments Used	Watkins-Farnum Performance Scale (Watkins & Farnum, 1954) AMMA (Gordon, 1989) Schematizing Test (Holzman, 1954) Kit of Factor-Referenced Cognitive Test (Ekstrom et al., 1976) Iowa Tests of Educational Development (Hoover et al., 2003)
Description	
Subjects	98 students from four public high schools
Instruments Used	Wind instruments
Stimuli	-
Apparatus	-
Method	Group tests: AMMA (for tonal and rhythmic audiation), Kit of Factor-Referenced Cognitive Tests (spatial orientation sub-tests: card rotations and cube comparison for mental rotation ability). After a break, or on another day: Schematizing Test for visual field articulation, and three subtests for spatial visualization from Kit of Factor-Referenced Cognitive Test Students were tested individually with WFPS
Results	Music sight-reading ability can be predicted by a combination of cognitive abilities: Reading comprehension, rhythmic sight-reading and audiation with aural feedback, spatial orientation
Reliability	-
Validity	-
Reference	Eastlund Gromko, J. (2004). Predictors of music sight-reading ability in high school wind players. *Journal of Research in Music Education, 52*(1), 6-15.

Title	A correlation study of keyboard sight-reading facility with previous training, note-reading, psychomotor, and memorization skills
Author	Jack Lyman Eaton
Year	1978
Purpose	• To analyze previous training or experience, note-reading, psychomotor, and memorization skills in relationship to keyboard sight-reading facility • To analyze sight-reading facility from the standpoint of the skills involved in successful development of that facility
Measurement	• Author-developed subject score sheet for note, rhythm, tempo errors Scoring provided by 3 faculty of the Jordan College of Music of Butler University 2 of these judges had masters' degrees in piano; the third judge was chairman of music theory and literature and a performing artist and instructor The judges scored independently, but at the same time, and discussed any discrepancies
Assessments Used	Subject data sheet for background musical information
Description	

Subjects	73 subjects, including university-level pianists and organists, and keyboardists from the community
Instruments Used	Piano
Stimuli	For sight-reading: A. Rigaudon (Daquin), B. Prelude Improvisation (Clementi), C. Silhouette (Dvorak), D. Preludio (Jarnach) For memorization skill: Minnelied (Hofmann), Study, op. 108, no. 2 (Schytte)
Apparatus	Poster boards for sight-reading test materials Index cars for note-reading, psychomotor, memorization skill materials *Panasonic* tape recorder for performances *Steinway Model "B"* grand piano for performances *Breitling* stopwatch for timing
Method	30-minute individual testing procedure The subject first completed the subject data sheet The order of testing instrument: Sight-reading facility, note-reading skill with verbal response and played response, psychomotor skill with visual command and oral command, memorization skill with playing practice and visual analysis
Results	Psychomotor skill achieved the highest correlation coefficient with sight-reading facility, followed by note-reading skill and years of keyboard playing experience
Reliability	Adjudicator reliability for scoring: Between .961-.995 for all combinations of adjudicators
Validity	-
Reference	Eaton, J. L. (1978). A correlation study of keyboard sight-reading facility with previous training, note-reading, psychomotor, and memorization skills (Doctoral dissertation, Indiana University). *Dissertation Abstracts International*, A 39 (07), 4109.

Title	The relationship among instrumental sight-reading ability and seven selected predictor variables
Author	Charles A. Elliott
Year	1982a
Purpose	• To investigate the relationships between instrumental sight-reading ability and seven selected variables • To determine which of the selected variables could, singly or in combination, account for variance in sight-reading scores
Measurement	• Selected variables: Technical proficiency, sight-singing ability, rhythm reading ability, cumulative GPA, cumulative music theory GPA, cumulative performance jury GPA, major instrument GPA Performance of WFPS and Criterion sightsinging test scored by two examiners each Technical proficiency test was scored by three examiners Examiners were graduate students in music theory or music education
Assessments Used	Watkins-Farnum Performance Scale (Form A) Author-composed technical proficiency evaluation Criterion sightsinging test (Thostenson, 1970) Rhythm reading test (Boyle, 1968)
Description	
Subjects	30 undergraduate students, selected at random from music theory classes
Instruments Used	Wind instruments
Stimuli	WFPS (Form A)
Apparatus	Tape-Recorder for all performances
Method	Subjects performed tests over a 6-week period
Results	Strong positive relationship between general sight-reading ability and ability to sight-read rhythm patterns Rhythm-reading ability is best predictor of sight-reading scores Rhythm-reading ability and performance jury scores are tougher the best predictors of

	sight-reading performance scores
Reliability	WFPS: .94 test reliability; rater reliability – 2 raters, .96 Technical Proficiency: Split-half correlation 098; rater reliability – 3 raters, .98 Criterion Sightsinging Test: .93 for pitch, .82 for rhythm, .92 for melodic, .95 for overall battery; rater reliability – 2 raters, .98 Rhythm Reading Test: Split-half correlation .92; rater reliability – 2 raters, .99
Validity	-
Reference	Elliott, C. A. (1982a). The relationships among instrumental sight-reading ability and seven selected predictor variables. *Journal of Research in Music Education, 30*(1), 5-14.

Title	The identification and classification of instrumental performance sight-reading errors
Author	Charles A. Elliott
Year	1982b
Purpose	• To identify and categorize sight-reading errors • To compare the types of errors made between “Superior sight-readers” and “poor sight-readers”; to determine if the differences are in number or in both number and kind
Measurement	• Errors being measured: 1. Pitch (key signature errors, missed accidentals, incorrect harmonics, other instances of incorrect pitch performed), 2. Rhythm (Pauses, hesitations, fluctuations of tempo, incorrect meter, incorrect interpretation of metronome marking, other instances of incorrect rhythm performed), 3. Expression (dynamic markings), 4. Articulation (slurring, tonguing, accent markings), 5. Other (errors that could not be categorized) Scoring procedure is outlined in the WFPS manual. Each error is marked. Tone and intonation are not considered. Only one error per measure is marked, and the total number of correctly performed measures comprises the score. Each performance was scored twice: First, independently by two instrumental music education graduate students, with a single sore averaged from these scores. The author also scored each performance, identifying every error committed and categorizing each. The author notes all errors
Assessments Used	Watkins-Farnum Performance Scale (Form A)
Description	
Subjects	Random sample of 30 undergraduate students, from music theory classes
Instruments Used	Wind instruments
Stimuli	WFPS (Form A) consists of 14 sight-reading exercises
Apparatus	Tape-recorder used for each performance Metronome for tempo
Method	All subjects receive Form A of the WFPS, proceeded with the tempo given by the metronome, and with a 15-20 s pause between each of the 14 exercises
Results	Each subject was assigned a sight-reading score to rank them: the upper 20% were “superior” and the lower 20% were “poor” Found that nearly two-thirds of total errors were rhythmic For poor readers, most errors were rhythmic. For superior readers, most errors were pitch Chi-Square value of 142.67, indication that there is likely a significant relationship between sight-reading ability and the types of errors committed
Reliability	Reliability of .94 reported for WFPS for grades 9-12
Validity	-
Reference	Elliott, C. A. (1982). The identification and classification of instrumental performance sight reading errors. *Journal of Band Research, 18*(1), 36-42.

Title	A pilot investigation of relationships between elementary keyboard sightreading achievement by music majors in college and selected musical profile tests
Author	Billie Erlings
Year	1977
Purpose	• To determine the possible predictive relationships between a specific, elementary level of sight reading achievement and two musical profile measuring instruments
Measurement	• Author-constructed measuring instrument for the evaluation of sight reading achievement quantitatively, at the required level for the university's keyboard proficiency requirement • Measuring pitch, rhythm, tempo, and continuity Scoring standards and procedures, developed from Dr. Erma Kleehammer's system Author-scored, and the two subtest scores were averaged. A second qualified piano teacher independently scored a sample of 1/3 of the tests
Assessments Used	Pretest: Identification of important variables in the sight reading materials, testing, and scoring procedures Experiment: Gordon's Musical Aptitude Profile (MAP): two sections used – Tonal Imagery, Rhythm Imagery Aliferis Music Achievement Test (MAT)
Description	
Subjects	33 non-keyboard music majors at university (21% vocal, 79% instrumental, 39% with no pre-college piano training); all enrolled in class piano. All had a minimum of previous piano training
Instruments Used	Piano
Stimuli	Gordon's Musical Aptitude Profile (MAP): Two sections used–Total imagery, Rhythm imagery Aliferis Music Achievement Test (MAT) Also used: GPA scores, American College test scores, and motor control was studied Author-constructed sight reading measurement: Two German dances (Haydn), arranged and transposed for control of variables
Apparatus	Computer to record keystrokes Goggles with a lower half of the frame painted black, and 2.5 inch black tubular extensions attached to lenses
Method	The two sight reading subtests (MAP and MAT) were given in two separate weeks, recorded on audio in a controlled setting
Results	Sight reading test scores range from 1-63.5 errors (0-34 was considered a passing score for the requirement). 67% were in the passing group. 38% of those with no precollege training passed Positive correlations found between sight reading and 1. MAP tonal imagery, rhythm imagery, composite, 2. MAP melodic section, rhythmic section, total score 1. Sections of the MAP and MAT proved to have a positive relationship with elementary keyboard sight reading analysis 2. The motor control test did not prove to be helpful. A more sophisticated eye-hand control test might be helpful
Reliability	-
Validity	-
Reference	Erlings, B. (1977). A Pilot investigation of relationships between elementary keyboard sightreading achievement by music majors in college and selective musical profile tests. *Bulletin of the Council for Research in Music Education*, 14-17.

F

Title	Dissociation in musical reading: A musician affected by alexia without agraphia
Author	A. M. Fasanaro, D. L. A. Spitaleri, R. Valiani, D. Grossi
Year	1990
Purpose	• To evaluate the functions of music reading and text reading in an alexic subject
Measurement	• Music-reading ability • Word-reading ability • Rhythms • Ideographic symbols • Notes
Assessments Used	***General Study*** *1. Intellectual level – Wechsler Adult Intelligence Scale, Raven's Progressive Matrices, 2. Memory – Immediate memory span – Corsi's Spatial Span Test, Wechsler Digit Forward Test, Learning – Spatial Learning on Corsi's Test, 4. Agnosia – Color agnosia - Ishihara Test, 6. Language examination – Understanding (Token Test)*
Description	
Subjects	72-year-old patient; professional violinist. He suddenly developed loss of strength and hypoesthesia. A neurologic exam showed disorientation, mild right hemiparesis with hypoesthesia, right hemianopia, alexia, and anomia. Further testing showed moderate left parieto-occipital slowing and an area of decreased density in the left temporoparieto-occipital region. The left medial region and splenium were damaged.
Instruments Used	Voice, Violin, Viola, Piano
Stimuli	***General Study*** *(See Assessments Used for specifics)* ***Reading*** *Lists of 110 words and 105 nonwords, and sentences. For matching test, 3-4-syllable words used as target words. On the possible matching list, there was the target word, a visually similar word, a semantically-related word, and a different word.* **Musical Abilities** Receptive Amusia – No specific stimuli listed Expressive Amusia – No specific stimuli listed Music Reading – Use of single notes, sequences of notes, rhythms, expressive symbols (e.g. crescendo, pp), other musical symbols (e.g. keys) Music Writing – No specific stimuli listed Matching Test – No specific stimuli listed
Apparatus	Blindfold, Pencil and paper
Method	3 stages of the Neuropsychological Study: 1. General neuropsychological aspects, 2. Reading tests, 3. Musical ability tests ***General Study*** *1. Intellectual level, 2. Memory – Immediate memory span, Learning, 3. Spatial exploring, 4. Agnosia – Color agnosia, Picture agnosia, 5. Apraxia, 6. Language examination, 7. Phonemic spelling (See Assessments Used for more detail)* ***Reading*** *Subject read aloud letters, words, nonwords, sentences in printed and script form* *Proprioceptive Reading – Subject was blindfolded and his finger was guided around the outline of letters, to transform the letter into proprioceptive information* *Matching Test – A word was given in another form (visual, auditory, proprioceptive) and subject selected written equivalent from a list* **Musical Abilities** Receptive Amusia – Subject was to identify pitches, melodies, and rhythms Expressive Amusia – Subject sang known melodies, ascending and descending scales,

	and performed key transposition. Also played these musical patterns Music Reading – Read single notes, notes in sequence, rhythms, expressive symbols, other musical symbols Music Writing – Dictation of single notes, dictation of musical symbols, dictation of rhythms. Copied sing notes, copied musical symbols, copied rhythms Matching Test – 5 types: 1. Named note matched with written note on staff. Chose from a list of 4 notes. 16 items, 2. Played note matched with written note on staff. Examiner played note on violin. 20 items, 3. Written name of a note matched with written note on staff. Subject shown written name of a note. 7 items, 4. Written note on staff matched with written note on staff. 20 items, 5. Played musical patterns matched with written musical patterns. Subject listened to brief piece of music played on violin, and identified corresponding piece from 4 scores written on staff. 10 items
Results	***General Study*** *Found memory impairment due to lesion in temporal lobe and thalamus of left hemisphere. Memory impairment prevented strategies to remember written letters. Affected by mild disturbance of visual analysis of the stimulus and showed mild anomic aphasia* ***Reading*** *Proprioceptive Reading – Only able to read letters and words with 5-6 letters (90% correct)* *Matching Test – Auditory word to written word: 80% correct, Proprioceptive word to written word: 90% correct, Picture to written word: 80% correct, Object to written word: 80% correct* *Found that reading was slow and letter-by-letter. Single letters were read better than words, and words were read better than nonwords. Subject presented letter-by-letter reading, or word form alexia. The phonological and global routes were impaired* **Musical Abilities** Receptive Amusia – Could identify the higher of two pitches. Was able to sing single pitches accompanied by piano. Able to identify known pieces of music, and to recognize intentional errors. Could identify instruments by their sounds. Could recognize rhythms Expressive Amusia – Could perform all tasks well. Could not play the musical patterns easily. Difficulty with identifying chords, placing hands in correct positions. Relative position of notes was always preserved Music Reading – Reading and playing single notes was 35% correct. Reading notes in sequence was 50% correct. Reading rhythms and other musical symbols was perfect Music Writing – Dictation and copying of rhythms and musical symbols was perfect. Dictation and copying of single notes was 70% correct. The errors were from incorrect positioning of the notes on the staff (a line or a space higher or lower than requested) Matching Test – Named note to written note, 100%, Played note to written note 65%, Written name of note to written note, 100%, Written note to written note, 90%, Played musical patterns to written musical patterns, 100% Found that subject non affected by amusia. Severe impairment with reading of notes, while reading of rhythms and other musical symbols was preserved. He recognized written pieces from their rhythms. The analytic and global routes were impaired; therefore subject had to evoke the internal representation by another route (auditory word, corresponding object, played note, named note)
Reliability	-
Validity	-
Reference	Fasanaro, A. M., Spitaleri, D. L. A., Valiani, R., Grossi, D. (1990). Dissociation in musical reading: A musician affected by alexia without agraphia. *Music Perception, 7*(3), 259-272.

Title	The effect of pattern recognition and tonal predictability on sight-singing ability
Author	Philip Fine, Anna Berry, Burton Rosner
Year	2006
Purpose	• To investigate the effects of disruption on sight-singing, with melody disruption, harmony disruption, and simultaneous melody and harmony disruption • To investigate the roles of pattern recognition and prediction in sight-singing performance • To investigate the roles of familiarity and experience in sight-singing ability
Measurement	• Interval-singing test for pattern recognition ability. Scored for accuracy and speed. An incorrect interval gained no marks • Sight-singing test scored for pitch only (not rhythm) and hesitation Two judges individually marked the productions of participants (One scored all and second judge scored a sample for reliability purposes)
Assessments Used	Author-composed interval-singing test Author-composed sight-singing test
Description	
Subjects	22 experienced choral singers, male and female students members of an accomplished chamber choir
Instruments Used	Voice
Stimuli	Interval-singing test: Pairs of notes instantiating 10 different intervals, from major third to major seventh, in ascending and descending form. These were divided into four sets, relating to SATB ranges of singers Sight-singing test: 4 complete Bach chorales were used as initial stimuli (BWV 6, 7, 74, 263). All 4/4 time, homophonic, simple rhythms, between 9-15 measures, all settings of unfamiliar German hymns. Each chorale was manipulated 3 ways: 1. Harmonic alteration for less tonal, more discordant, and less predictable, 2. Melodic alteration for more random and irregular with increased interval sizes, 3. Harmonic and melodic alteration to disrupt melody and harmony, with only rhythmic structure of original chorale intact. Total of 40 stimuli, including the different voice parts and alterations
Apparatus	*Clavinova* digital piano for recording of stimuli Cassette tapes to play for participants, with stimuli and verbal instructions Cassette player to record performances of participants
Method	Subjects first completed the interval-singing test, including 10 written pairs of notes. The first note was played from the tape and the subject was to sing the interval as quickly and accurately as possible. For sight-singing test, the participants sang the four chorales slowly, with a second attempt at each. A measure of beats on the starting pitch preceded each chorale
Results	Interval-singing performance correlated with overall sight-singing performance, suggesting that skilled sight-singers are better at recognizing and using melodic patterns and these develop with experience. Altered melody impaired performance. Altered underlying harmony impaired ability, significantly more so with less-skilled readers. Notational audiation (also called inner hearing) is important in instrumental sight-reading and sight-singing. For singers, auditory feedback helps to keep them in tune. Concludes that pattern recognition and prediction are important abilities
Reliability	Inter-judge reliability: .84
Validity	-
Reference	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.

Title	The effects of skill on the eye-hand span during musical sight-reading
Author	S. Furneaux and M. F. Land
Year	1999
Purpose	• To examine the eye-hand span (EHS), measured in notes and in time
Measurement	• Eye movements: each fixation on score, fixation duration, saccadic amplitude and duration, timing of movements • EHS in time – The length of time between fixation and performance; the 'time index' • EHS in notes – The number of notes between hand and eye; the 'note index' Above measured for pianists at different skill levels • Effect of tempo on EHS
Assessments Used	-
Description	
Subjects	8 adult pianists: 3 'novices' (ABRSM Grade 3-4), 3 'intermediates' (ABRSM Grade 6-7), 2 professionals (accompanists)
Instruments Used	Piano
Stimuli	5 pieces of music. Each piece was assigned 2 tempos: "very slow" and "very fast". The pieces were short extracts of musical scores published under a particular grade standard.
Apparatus	Head-mounted video camera system (Land 1993) Two mirrors for split-screen video recording Metronome to introduce tempo
Method	Subjects were calibrated with head-mounted video camera. They were to sight-read as soon as stimuli were presented.
Results	Found that pianists read the 2 lines of information on dual-staved music separately, with a zigzag pattern of eye movements and fixations. Principal findings: 1. The note index increases with an increase in skill (suggests that buffer increases, so more proficient musicians are able to store more information), 2. The mean time index does not change with skill (average of 1 s for all subjects), 3. The time index reduces with an increase in performance tempo for all skill levels.
Reliability	-
Validity	-
Reference	Furneaux, S. & Land, M. F. (1999). The effects of skill on the eye-hand span during musical sight-reading. *Biological Sciences, 266*(1436), 2435-2440.

G

Title	Restricting the field of view to investigate the perceptual spans of pianists
Author	Elizabeth Gilman and Geoffrey Underwood
Year	2003
Purpose	• To determine how the perceptual span of pianists varies with developing skill and cognitive load • To investigate the perceptual span of pianists while they read more complex music phrases in order to establish any differences in perceptual span between good and poor sight-readers
Measurement	• Eye-movements through a gaze-contingent window, with 1, 2, or 4 beats (sampled every 4 ms) • Fixation duration • Note duration • Position of first error • Number of errors • To provide an accurate measure of the EHS during sight-reading and transposition, by measuring the location of the point of fixation at the onset of each note/chord and calculating the horizontal difference between the point of fixation and the point of performance. The EHS (i.e., how far the musician is reading ahead of the point of performance) can then be combined with the perceptual span (i.e., the amount of useful information extracted around the point of fixation) to give an overall measure of how much useful information is extracted ahead of the point of performance (Truitt et al., 1997) P. 206-207 • Finally, we measured the effect of sight-reading skill on eye-movement and performance measures, and on EHS and perceptual span, by comparing good and poor sight-readers of matched general musical ability
Assessments Used	-
Description	
Subjects	40 subjects, all who had achieved grade 8 piano (in UK), or equivalent. Subjects were matched on general performance ability and allocated to 2 sight-reading groups on basis of pre-test. Those who scored higher than median were "good sight-readers and those who scored below were "poor sight-readers"
Instruments Used	Piano
Stimuli	32 musical phrases from Bach chorales (Riemenschneider, 1941 (Ed.)). Tenor parts were removed. All phrases 3 bars long with 3-beat rest at beginning and 1-beat rest at end. These were used for sight-reading and transposition tasks. For error detection task, phrases were further edited to have one-four notational "errors". Errors made by shifting one-four notes either up or down to have music no longer make harmonic or melodic "sense"
Apparatus	*CuBase Score* for stimuli creation *GraphicConverter* for stimuli presentation *SMI Eyelink* eye-tracker *Compaq Deskpro* computers for presentation and eye-tracker control 17-inch monitor for viewing *C++* program for moving window *Casio 3800* keyboard connected to MIDI interface
Method	Subjects seated 1 m from monitor. They rested their head on a chin support. Eye-tracker was calibrated. Each subject performed three tasks in randomly determined order: Sight-reading task, transposition task, error-detection task. Sight-reading: 32 phrases in random order, and subject performed each phrase at a comfortable tempo. Transposition Task: Same 32 phrases presented in random order, and subject

	performed phrase 2 semitones below the notes indicated in score. Error-detection task: 32 edited phrases presented, and subjects knew that phrases would have at least one notational error. They searched through phrases as if sight-reading and indicated how many errors they thought each phrase contained using numbers on computer keyboard. For each task, subjects viewed ¼ of phrases through 1-beat window, ¼ phrases through 4-beat moving window, and ¼ of phrases with no window.
Results	Found that good and poor readers do not differ with perceptual span, but good readers found to have larger eye-hand spans. Also found that increasing cognitive load decreases eye-hand span, but has little effect on perceptual span
Reliability	-
Validity	-
Reference	Gilman, E., & Underwood, G. (2003). Restricting the field of view to investigate the perceptual spans of pianists. *Visual Cognition, 10*(2), 201-232.

Title	Towards an eye-movement model of music sight-reading
Author	Elizabeth R. Gilman
Year	2000
(General) Purpose	• To test whether the Kinsler and Carpenter (1995) model of eye movements during sight reading can account for processing of pitch and harmony information, focusing on the behavior of pianists. Also aims to test the proposed effect of complexity, predictability and eccentricity, using 10 tasks that eliminate confounding effects of performance Considers an alternative model of sight-reading called the “EZ Reader”, designed for text reading. Limited to the reading homophonic styles of music; in particular, the reading of chordal structures. Experiments are therefore designed to find out how chords are processed. Experiments are designed to investigate the effects of tonal complexity on processing, and the differential effects of complexity for sight-readers of varying levels of skill. • Thesis also examines 2 alternative measures of predictability: context-based predictions and associative priming. It investigates their effect on the processing of musical notation • Also attempts to find out whether or not the processing of parafoveally-presented notation is an all-or-none process • Finally, thesis investigates effects of predictability and eccentricity on eye-movement behavior in order to find out whether or not these two variables interact.
Pre-Test: Sight Reading	
Measurement	• Sight reading scored by quantizing the performance, with durations of notes and rests rounded to nearest semiquaver beat. Notes played that corresponded most highly with notes that should have been played were chosen on basis of 1. Overall position in the score, 2. Pitch, 3. Temporal value. Rests were chosen on basis of their position and their temporal value. Pitch and rhythm scores coded separately. Pitch score = 1 point for each note played. 1/12 of a point deducted for each semitone between the note played and corresponding note in the score. Rhythm score = 1 point for each note played and for each rest accounted for. 1/4 of a point deducted for each semiquaver beat between the duration of the played note/rest and duration of the corresponding note/rest in the score. Durations of extras notes and rests were added together and 1/4 of a point was deducted from rhythm score for each semiquaver beat. Finally, pitch and rhythm scores were added together and divided by the highest possible score to calculate a percentage.

	Scores were compared with an expert's evaluation of the performances. This examiner rated technical ability (out of 40) and expression (out of 20).
Assessments Used	-
Description	
Subjects	16 grade 8 pianists, 5 grade 6 pianists, 3 grade 5 pianists
Instruments Used	Piano
Stimuli	Subjects performed at least 1 specified piece of music; 2 at most. They were grade 8 level from Guildhall School Graded Piano Sight-Reading series. Slow piece was "Lament" in 4/4 time and C minor. The faster piece was "Habanera" in 4/4 time and G minor. Both pieces altered slightly
Apparatus (for this and other experiments)	*CuBase Score V. 1.1.1 and V. 4.1* recorded and then created visual representation of performance *(Used for any experiments with keyboard response)* Metronome for continuous tempo during performance *Yamaha YPP-35* touch-sensitive keyboard for performance *GraphicConverter* to edit and convert MIDI data into a score using PICT format *Psyscope V. 1.2.1* used to present stimuli and record responses *(For experiments 1-8)* *Dual purkinje image eye-tracker* for measuring eye-movements *(experiments 9-10)*
Method	Subjects performed the 2 pieces, attempting the slower piece first. They had 2 minutes to look at each piece before playing
Results	Subjects were divided into "good" or "poor" readers, depending on whether they achieved higher or lower than 75%
Reliability	Sight-reading scores on objective measure with expert's marks r = 0.876
Validity	-
Experiment 1: Chord Naming	
Purpose	• To investigate the processing of chordal structures by good and poor sight-readers. To find out whether chords are likely to constitute the 'chunks' of homophonic music reading. Is chord recognition sensitive to expertise effects?
Measurement	• Time taken from start to finish for naming each chord
Assessments Used	-
Description	
Subjects	40 pianists; all having passed the grade 8 piano exam or foreign equivalent. 20 of these scored 75%+ on pre test and were considered "good sight-readers". Those who scored below 75% were considered "poor sight-readers"
Instruments Used	-
Stimuli	72 chords created: 12 major, 12 minor. These were presented in 3 inversions: root position, 1st inversion, 2nd inversion. They were all presented as semibreves on staves with appropriate clefs. All notes were presented with accidentals
Apparatus	*CuBase Score 1.1.1* for chords 17" monitor for presentation *Psyscope* on a *Macintosh Quadra 900*
Method	Subjects sat at computer monitor. They fixated on cross in centre of screen, and then chord was presented and remained until subjects named the chord. Subjects were required to name chord by identifying the tonic note and the tonality.
Results	Found significant difference in time taken to start naming chords between the two groups. Good readers are significantly faster and more accurate at naming chords. Suggests that good readers read music more effectively and efficiently because they identify chords as whole units rather than series of individual notes. Found that subjects responded faster to chords learnt earlier in their career.
Reliability	-
Validity	-
Experiment 2: Processing Chords and Dischords	
Purpose	• To investigate the processing of chordal structures by good and poor sight-readers. To test for skill differences when musicians are presented with chords

	and discords
Measurement	• To measure the processing of notational information from one fixation
Assessments Used	-
Description	
Subjects	The 40 pianists from Experiment 1
Instruments Used	
Stimuli	36 majors chords and 36 minor chords created. They were in root, 1st and 2nd positions. All had 4 notes, with accidentals. Then, 36 discords were created to visually match the major and minor chords; Their 4 components did not all belong to same key.
Apparatus	*CuBase Score* for chords 17" monitor for presentation *Psyscope* on *Mactintosh Quadra 900*
Method	Subjects were seated at computer screen and presented with chords and discords for 180 ms. They were to immediately recall as many notes as possible, and draw the 4 notes on an open stave, then press mouse key for next trial. The subjects were always told to write 4 notes, even if they were guessing
Results	Found that good sight readers outperformed poor sight readers. Found that poor sight readers responded to major/minor chords more accurately than to discords. Suggests that poor readers (and all readers) do not process chords note-by-note but identify patterns of notes that form familiar structure, and process the constituent notes in parallel. Good readers create and store more chunks than less skilled readers, including chunks for late-acquired musical patterns
Reliability	-
Validity	-
Experiment 3: Transposition Paradigm	
Purpose	• To compare the transposition performance of good and poor sight readers with tonal and atonal stimuli • To develop a visual analogue of tasks that have used transposition in order to study interval-processing • To find out if the transposition task is sensitive to skill differences and to discover whether any effect of skill is influenced by the musical style of the stimuli
Measurement	• The processing of intervals between notes • Mean response time to each trial • Number of correct trials
Assessments Used	-
Description	
Subjects	8 good sight readers (scored 70%-100% on pre test) and 8 poor readers. There was also a group of 8 'less-skilled' readers; pianists who had achieved Associated Board grade 6 or 5 piano. Note that poor and less-skilled readers did not differ in terms of sight reading ability
Instruments Used	-
Stimuli	64 stimuli (4 sets of 16); each 2 lines of melody, presented directly below each other. Set A: 4-bar melody on top line, called the 'original', and transposition on bottom line. This correct transposition is composed of the same intervals from the original melody. Set B: Original on top line and transposition on bottom line, but with one of the notes as an incorrect transposition. Set A and B conformed to traditional harmony rules (excluding transpositional errors) and considered 'tonal'. Presented with conventional music notation with key signature. Set C: Original on top line and correct transposition on bottom line. Set D: Original on top line and transposition with one error on bottom line. Sets C and D were composed in modern style which does not conform to traditional harmony rules. Considered 'atonal'. These were written without key signatures, but used accidentals. For all stimuli: 13 notes in each melody line, half written in treble clef and half in

	bass clef, half were in 3/4 time and half in 4/4 time
Apparatus	*Psyscope* on *Macintosh Quadra 900* to present stimuli 17" computer screen
Method	Based on the "pattern-matching" task (See Waters & Underwood, 1998). Subjects are exposed to melody, followed by a correct or incorrect transposition of the melody. Beginning of each trial indicated by cross presented in centre of screen for 500 ms. Immediately follow by one of 64 stimuli. Stimulus presentation terminated when subject pressed one of 2 buttons: yes or no. Subjects were to respond as quickly as possible. 1200 ms interval of blank screen between trials
Results	Found that subjects took longer to respond to atonal stimuli. Good readers performed transposition task more accurately than the other 2 groups. This indicates that processing of intervals is important cognitive sub-component of skill reading. Poor readers outperformed less skill readers. Poor and less skilled readers diff in terms of general musical ability rather than sight reading ability. Either, all 3 groups make use of a visual coding process and a more abstract encoding process. This abstract process is faster and more developed in good reader than in other groups. OR, all readers use only an abstract coding mechanism, but that it is attuned to more complex representations in good readers.
Reliability	-
Validity	-
Experiment 4: Interval-Matching Paradigm	
Purpose	• To directly compare the processing of visually similar and visually dissimilar intervals by good and poor sight readers in an attempt to distinguish between visual and abstract methods of interval encoding • To test 2 alternative models of interval encoding suggested in Experiment 3
Measurement	• Distinguish between visual and abstract methods of interval encoding
Assessments Used	-
Description	
Subjects	The 40 pianists from Experiments 1 and 2
Instruments Used	-
Stimuli	10 different target intervals created. They were 2 notes presented on a 5-line musical stave with treble clef. For each target, a visually similar/same interval and a visually similar/different interval created. Visually similar = 2 notes separated by the same number of lines and spaces as target. Accidentals, if necessary, would be in the same positions. Visually similar/same = 2 constituent notes separated by same number of semitones as target. Visually similar/different = 2 constituent notes separated by different number of semitones as 2 notes in target. Then, for each target, a visually different/same interval and visually different/different interval created. Visually different = 2 notes separated by same number of lines and spaces as target. If target had accidentals, visually different intervals annotated with different accidentals or no accidentals. If target did not have accidentals, visually different would contain an accidental. Visually different/same = 2 constituent notes separated by same number of semitones as 2 notes in target. Visually different/different = 2 constituent notes separated by a different number of semitones as 2 notes in target. For each target, spatially different/same interval and spatially different/different interval created. Spatially different = 2 notes separated by a different number of lines and spaces than target. Spatially different/same = 2 constituent notes separated by same number of semitones as target. Spatially different/different = 2 constituent notes separated b a different number of semitones as target. 12 practice stimuli also created
Apparatus	*Psyscope* on *Macintosh Quadra 900* to present stimuli 17" computer screen for presentation
Method	Subjects presented with practice block of 6 trials, followed by 2 blocks of 30 random,

	counterbalanced trials. Each trial began with fixation cross for 500 ms. Comparison interval presented and subjects were to calculate interval between 2 notes, then press mouse button. Then target interval presented. For 1/3 trials, comparison was visually similar to target. For 1/3 trials, comparison was visually different from target. For 1/3 trials, comparison interval was spatially different from target. Subjects calculated interval and indicated whether target was same or different to comparison interval by pressing computer key. Half of the trials were same and half the trials were different.
Results	Results are in favour of a dual coding model of interval coding. Poor readers are more vulnerable to the modulating effects of visuo-spatial features than are good readers. Establishes that abstract processing mechanism is less developed in poor readers and therefore they tend to rely on processing of visuo-spatial features.
Reliability	-
Validity	-
Experiment 5: Context Effects on Sight-Reading Chords	
Purpose	• To study the nature of context-based predictions formed when sight reading music. Attempts to identify the stage of processing at which expectation effects occur. Aims to empirically test the effects of context on good readers and poor readers
Measurement	• Time taken to perform the target Analysis only performed on scores from the 24 subjects who had rate of 60%+
Assessments Used	-
Description	
Subjects	40 pianists; all having passed grade 8 piano exam or foreign equivalent. Divided into 20 good readers and 20 poor readers. 40 subjects divided into 4 equal-sized groups: 1, 2, 3, and 4.
Instruments Used	Piano
Stimuli	Created from 2-bar phrases from Bach Chorales. All phrases in homophonic style and ended in cadence. Phrases in 4/4 time and began on 4th beat of the bar. Tenor part was removed from original chorale. The final chord and rest labeled as 'target' while rest of phrase labeled as 'context'. Phrases paired so that context of one phrase was incongruent with target of its pair. Therefore, each phrase had a congruent context and incongruent context. 40 targets in total. Divided into 4 groups: Group A phrases paired with Group B, and the same for Groups C and D. On half of the trials the notes were in black and on other half they were in light grey (degraded).
Apparatus	*Psyscope* on *Macintosh Quadra 900* to present stimuli 17" computer screen to present stimuli Metronome
Method	Each subject seated at computer screen and saw 40 trials. At the beginning of each trial, a context appeared and metronome beat was initiated. Subject listened to metronome for 3 beats, then performed the musical phrase. As they performed the penultimate chord of the phrase, the final chord appeared, after a delay of 500 ms. They were required to play this chord as quickly as possible.
Results	Results show that sight-readers show evidence of context effects while reading music. They perform target significantly more rapidly and accurately when target is preceded by congruent context. Support previous findings that good readers form context-based expectations and poor reader form fewer expectations; therefore less likely to be affected by context. Good readers consistently accurate when performing target chord, even when context was incongruent. Good readers only affected by context in terms of response latency.
Reliability	-
Validity	-
Experiment 6a: Triad Priming	
Purpose	• To investigate the role of 'harmonic priming' during music reading
Measurement	• Sight-reading skill • Target triad stage of acquisition (StOA)(early-acquired targets and late-acquired

	targets) • Relatedness between prime and target triads • Trial type
Assessments Used	-
Description	
Subjects	40 pianists from Experiments 1, 2, 4
Instruments Used	-
Stimuli	48 triads used as priming stimuli. These comprised tonic root position of all possible major and minor keys. Each triad made up of tonic note, with median and dominant notes above. Each was a quarter note triad on a 5-line staff, with treble clef. Accidentals used as needed. These same 48 triads used as target, but annotated on staff without treble clef. Early-acquisition triads: C, G, D, A, E, F major; A, B, E, G, D, C minor. All others were late-acquisition triads. For each target, 4 possible primes: related major, related minor, unrelated major, unrelated minor
Apparatus	*CuBase Score* to compose triads *GraphicConverter* for editing triads Slide with fixation cross Macintosh Quadra 900 for stimulus display 17" monitor for stimulus display
Method	There were 96 trials. At the beginning of each, a fixation cross was in centre of screen for 500 ms. Then subjects saw priming triad. After 1000 ms target triad presented to the right of the prime. Both triads remained on screen until response made. Subjects were to decide whether first triad presented was a major or minor triad, than to decide whether second triad presented was major or minor. If subject thought both were major/minor, they responded with "same" key. Otherwise, they responded with "diff" key.
Results	Good readers performed the task significantly more rapidly. Good readers significantly more accurate. Results appear to display evidence of priming effect: Musicians significantly more accurate and faster (thought not significantly faster) when presented with related prime-target pairs than when presented with unrelated pairs. Significant interaction between effect of relatedness with stage of acquisition of target chord: Early-acquired chord had priming effects. Late-acquired target had no priming effect. "At best, we cannot tell from the results of this experiment whether triad priming is an existing phenomenon." (Gilman, 2000, P. 183).
Reliability	-
Validity	-
Experiment 6b: Chord Priming	
Purpose	• To investigate the role of 'harmonic priming' during music reading
Measurement	• Sight-reading skill • Target chord stage of acquisition (StOA) • Relatedness between prime and target chords • Trial type
Assessments Used	-
Description	
Subjects	The same 40 subjects from Experiment 6a (and Experiments 1, 2, 4)
Instruments Used	-
Stimuli	98 prime-target pairs created. Natural chords used as stimuli. Related prime-target pairs described as "associated" and possible in Western tonal music. Unrelated prime-target pairs described as "unassociated". Prime and targets were 4 notes – 2 in treble and 2 in bass. All were semibreve on 5-line staff. Accidentals used. Priming triads had clefs; targets did not. Root position tonic chords of all 24 major and minor keys were targets. Each had 4 different paired primes: Related major, related minor, unrelated major and unrelated

	minor. All prime-target pairs composed with constraint that target chord was msical plausible. This meant that melody notes were never more than tone apart, and it would be possible for pianists to play in succession, and some contrary motion in the clefs. Early- and late-acquired chords same distinction as above experiment.
Apparatus	*CuBase Score* for creating stimuli *GraphicConverter* for editing stimuli Slide with fixation cross *Macintosh Quadra 900* for presentation 17" monitor for presentation
Method	96 randomized trials. Subjects saw fixation cross for 500 ms, then priming chord. After 1000 ms the target chord presented to the right of the prime. Both chords remained until subject made response. They were to decide whether first chord presented major or minor chord, then whether the second chord was major or minor. If they thought they were the same, they pressed "same", If they thought that one was different, they pressed "diff"
Results	Found that good readers performed task significantly faster, and made fewer errors. Related pairs responded to more rapidly. Musicians responded more rapidly to early-acquired targets. Major-major pairs responded to faster. Good and poor readers were affected by StOA, but effect was much stronger for poor readers. Suggests that good readers' chunking mechanisms are sophisticated enough to chunk both early- and late-acquired chords.
Reliability	-
Validity	-
Experiment 7a: Developing a New Priming Paradigm	
Purpose	• To implement serial rather than parallel priming in order to eliminate the stage of acquisition artifact • To increase ecological validity by measuring responses that involve performing a target chord • To include a stimulus onset asynchrony factor and a neutral condition so that conscious and unconscious process can be identified.
Measurement	• Relatedness and stimulus onset asynchrony (SOA) 3 levels of relatedness: Related, unrelated and neutral 2 levels of SOA: Short SOA (650 ms) and long (2500 ms) • Time taken to perform a target triad • Percentage of target triads performed correctly
Assessments Used	-
Description	
Subjects	The 40 subjects from Experiments 6a and b (and other experiments)
Instruments Used	
Stimuli	Similar to stimuli in Experiment 6a. 48 triads used at both priming and target stimuli. All triads were annotated by quarter note on a 5-line staff with treble clef. Accidentals used. Additionally, a "rest prime" (that is, a quarter rest) stimulus and a mask (looks like a large group of notes on top of each other) stimulus were created. On 1/3 of trials, target was related minor. On 1/3, unrelated major. On 1/3, unrelated minor. 4 trial types: Major-major, major-minor, minor-major, minor-minor. Also 24 neutral trials with "rest prime"
Apparatus	*Psyscope* for stimulus 17" monitor for presentation *Macintosh Quadra 900* for presentation *CuBase Score 1.1.1* for recording
Method	Subjects saw 2 blocks of 120 trials in counterbalanced order: Block A and B. Each trial began with fixation cross for 500 ms. Then presentation of prime stimulus for 600 ms (Block A) and 2450 ms (B). Subject looked at prime but did not perform this prime. After 50 ms, target stimulus presented in place of the prime. Subject was to perform the second stimulus as quickly as possible. Target remained until triad was

	performed, followed by a mask for 2000 ms to indicate end of trial
Results	Good readers consistently outperformed poor readers. Found no evidence of any priming effects. This is likely because subjects were not fully processing the prime, although they must have been processed to a certain extent, even if only the visual features. No significant differences found between response times to targets with related and unrelated primes
Reliability	-
Validity	-
Experiment 7b: A New Priming Paradigm	
Purpose	• To determine if the stage of acquisition of the prime has an effect on the processing of the target
Measurement	• Relatedness, StOA, and trial type 2 levels of relatedness: Related and unrelated 2 levels of StOA: Early-acquired and late-acquired 4 levels of trials type: Major-major, major-minor, minor-major, minor-minor • Time taken to perform a target triad • Percentage of target triads performed correctly
Assessments Used	-
Description	
Subjects	40 pianists with grade 8 piano (Associated Board exam). 20 pianists scoring 75%+ on pretest were “good” and others were “poor” readers
Instruments Used	Piano
Stimuli	Stimuli identical to those used in Experiment 7a. Also included “dummy” trials, with a question mark instead of a target, to indicate that the subject should perform the prime rather than the target triad. This was to ensure subjects processed the prime stimulus and that they had to retain the prime stimulus in their memory
Apparatus	*CuBase Score* for recording *Macintosh Quadra 900*
Method	Subjects saw 144 trials. At the beginning of each trial, fixation cross for 500 ms. Then a prime triad for 600 ms. On 96 trials, this was followed by target triad, presented until subject performed the triad. Half of these were ‘related’ and half were ‘unrelated’. On the remaining 48 trials, prime was followed by a question mark. If this was presented, they had to perform the triad they had just seen. A mask was presented to indicate end of trial
Results	Good readers outperformed poor readers. Subjects responded more rapidly when targets were early-acquired than late-acquired. Evident of a major superiority effect where major targets were responded to more rapidly than minor targets. Found no significance, but that the interaction shows that early-acquired triads are sensitive to priming effects whereas no evidence to suggest that late-acquired triads are sensitive to same effects. In summary, there is clear evidence of triad-priming, although priming effects are limited to early-acquired targets
Reliability	-
Validity	-
Experiment 8: Processing Extrafoveally Presented Chords	
Purpose	• To compare how well the Kinsler & Carpenter model and the EZ reader model account for the effects of eccentricity on the processing of music notation • To study the limitations of processing extrafoveal notation, namely triads. In other words, the experiments measure how processing declines with eccentricity • To investigate the nature of information obtained from notation presented extrafoveally. Specifically, it investigates whether the musician gains preview of the notes about to be fixated, or whether the musician just gets some idea about the contour of the notation
Measurement	• Extrafoveal processing

	2 levels of task: 'inversion' task and 'harmony' task 2 levels of direction: 'left' and 'right' 3 levels of eccentricity: 1.9", 3.8" (Approx. 1 bar of music), 5.7" visual angle
Assessments Used	-
Description	
Subjects	40 subjects from Experiments 1, 2, 4, 6, 7a
Instruments Used	-
Stimuli	6 root position triads: C, F, G Major; d, e, a minor. Also composed triads in 1st and 2nd inversion with same tonics and tonalities Inversion task: Each of 3 possible inversions presented in centre of screen a total of 24 times. Each stimulus was an inversion of one of the 6 triad types. For each type of inversion, an "extrafoveal triad" was presented 4 times at each of the 6 extrafoveal locations. Therefore, total of 72 trials presented in randomized order
Apparatus	*CuBase Score* for stimuli *Psyscope* on *Macintosh Quadra 900* for presentation 17" screen for viewing
Method	Subjects sat at computer screen. Each trial began with fixation cross for 500 ms. Then a triad in the centre of the screen and a different triad presented either 1.9", 3.8", or 5.7" to the right or left of central triad. Both displayed simultaneously for 300 ms. Subjects were to assume that all of these were in treble clef and without a key signature. Inversion task: Subjects stared at centre of screen for each presentation, then pressed a button to indicate whether triads were the same inversion or different inversions, regardless of harmonies. Harmony task: Subjects were to indicate whether triads had same tonic note, regardless of inversion. Then subjects annotated on a blank stave the triad they had seen in the centre of the screen
Results	Good readers significantly outperformed poor readers. Supports hypothesis that good readers can encode musical notation presented extrafoveally, and a reason why they sight-read more efficiently. The inversion task was much easier than the harmony task. Readers performed more accurately when extrafoveal chords were presented to right of fixation then when chords were presented on the left. This suggests left hemisphere is more important then right hemisphere for recognition of musical notation. Good readers outperformed poor readers only when extrafoveal chords were presented at 1.9" and subjects were matching on basis of harmony. Otherwise, no significant effect. Suggests that good readers have greater note identification spans that poor readers, but no difference between their perceptual spans. Likely conclusion is that good readers do in fact have larger perceptual spans than poor readers, but that the effect of skill on perceptual span is not as great as effect of skill on note identification span. In other words, perceptual span of good reader extends beyond their note identification span
Reliability	-
Validity	-
Experiment 9: Extrafoveal Processing with Eye-tracking	
Purpose	• To study the limitations of processing extrafoveal notation, namely triads. In other words, the experiments measure how processing declines with eccentricity • To attempt to replicate the findings of Experiment 8 using eye-tracking equipment to ensure that the subjects fixate the central triad
Measurement	• Eye-movements, while subjects are seated farther away from screen. Eye-movements sampled every millisecond 2 levels of task: 'Inversion' and 'Harmony' 2 levels of direction: 'Left' and 'Right' 3 levels of eccentricity: 1.4", 2.9", 4.3" visual angle

Assessments Used	-
Description	
Subjects	24 pianists; all having passed Grade 8 exam (Associated Board) or foreign equivalent. 4 subjects were eliminated from experiment. Of remaining 20, 10 were considered 'good' (scored 75%+) and 10 considered 'poor'
Instruments Used	-
Stimuli	Triads composed: C, F, G Major; d, e, a minor; b diminished. 3 different inversions of each triad were created. Each of above triads (except a minor and b diminished) were presented as central triad a total of 12 times. Total of 180 stimuli into 3 blocks: A, B, C. In each block, each triad presented 12 times as central triad: twice in each inversion with an extrafoveal triad in same inversion, and twice in each inversion with extrafoveal triad of same tonic note.
Apparatus	*CuBase Score* for stimuli 14" monitor for viewing Eye-tracker with chin rest and *Velcro* straps
Method	Subject seated 80 cm from screen and rested chin on rest provided. Velcro straps fastened around subject's head. Subject was to keep as still as possible throughout experiment. Then eye-tracker was calibrated. Subjects performed an inversion task and harmony task. Same stimuli presented in both tasks, but with different instructions. For each trial, stimulus presented for 180 ms. Inversion task: Subjects pressed 'yes' is both triads were written in same invesion, and 'no' if triads were not in same inversion. Harmony task: Subjects pressed 'yes' if triads had same tonic, and 'no' if they had different tonic notes. Subjects were to stare at central triad at all times, with 2 min break between each block
Results	Results similar to Experiment 8. Good readers performed significantly more accurately than poor readers. Readers were more accurate with inversion task. Performance declined with greater eccentricity. Suggest that good readers are more proficient at reading musical notation presented extrafoveally. Readers more proficient at inversion task. In conclusion, results favour the EZ Reader model.
Reliability	-
Validity	-
Experiment 10: Using a Matching Task to Investigate the Effects of Expectation and Preview	
Purpose	• To investigate the nature of information obtained from extrafoveally presented musical stimuli • To investigate whether the effects of expectation and parafoveal preview interact when musicians sight-read music. In other words, do expectations mediate the use of parafoveal preview when sight-reading? • Also investigate whether these effects can be explained within the framework of the Kinsler and Carpenter model of sight-reading and the EZ Reader model of text reading • To investigate the interaction between context and preview in the domain of sight-reading
Measurement	• Eye-movements (sampled every 3 ms) • Effects of expectancy and preview during sight-reading 2 levels of expectancy: "Expected" and "Unexpected" 2 levels of preview: "Proximal" and "Distal" • Score • Mean fixation duration • Number of fixations
Assessments Used	-
Description	
Subjects	8 'good' readers and 8 'poor' readers, who also took part in Experiment 7b (Only 16 of original 40 subjects participated due to vision). Equal number of subjects allocated

	to each of 4 groups: 1, 2, 3, 4
Instruments Used	-
Stimuli	Phrases created from 56 4-bar phrases from Bach's chorales (Riemenschneider, 1941). Written on grand staff, ended in cadence, and no key changes. Phrases edited by experimenter so that all comprised 35 notes in total. Phrases divided into 4 groups: A, B, C, D. These were matched on number of sharps or flats in key signatures. 4 version of each phrase created: "expected proximal" (EP), "expected distal" (ED), "unexpected proximal" (UP), "unexpected distal" (UD). Expected versions had original notes of phrase, and unexpected versions were created by shifting notes in upper or lower stave up or down by one stave position. This cause the notes of the phrase to no longer make musical sense, while contour of phrase and number of accidentals remained intact. Proximal stimuli were closer together (distance between notes = 11 mm) and distal stimuli were stretched out (distance between notes = 22 mm); all stimuli was three measures per line. For each of the 224 phrases, a "nonidentical twin" created by shifting one note in the top stave up or down by one line and one space, and shifting one note in bottom stave by same, in opposite direction
Apparatus	*Dual Purkinje Image* eye-tracker *CuBase Score* for stimuli 14" monitor for presentation
Method	(Based on matching task developed by Waters, Underwood & Findlay (1997) and Waters & Underwood (1998)). Subjects seated 1 m from display monitor. They were connected and calibrated to eye-tracker. Each subject saw 56 trials in total. Subjects were presented with randomly selected phrase from their assigned group and asked to read through phrase as if they were sight-reading. They pressed mouse button when at end of phrase. Then "nonidentical twin" phrase was presented and subjects instructed to read through until they found difference between this and previous phrase. When they found the note that differed, they were asked to stare at the note while pressing mouse button
Results	Found that musicians read music using fewer, shorter fixation when notes in the score can be predicted from surrounding context. When musicians are required to perform predictable music, saccadic frequency and saccadic amplitude are both decreased. Musicians used fewer fixations to read the scores when notes were proximal, suggesting that they are more likely to skip notes if parafoveal preview is increased. Fixations were longer in proximal condition. Accuracy data indicated that musicians were superior at detecting differences between phrases in expected condition. Performance was superior in the proximal condition. In summary, when music was expected, readers used fewer, shorter fixations. When music was 'proximal', reader used fewer, but longer, fixations. Effects of expectation and preview when sight-reading are additive rather than interactive (however, possible due to manipulations of this specific experiment)
Reliability	-
Validity	-
General Discussion	The Kinsler and Carpenter model of sight-reading is insufficient. The EZ Reader appears to be an appropriate model to explain music sight-reading, although I may require further research into the interaction between the composite variables. This model needs to be further modified to explain expertise differences
Reference	Gilman, E. R. (2000). *Towards an eye-movement model of music sight-reading* (Volume 1 and 2). University of Nottingham, Nottingham, United Kingdom.

Title	Eye movement in music reading: Effects of reading ability, notational complexity, and encounters
Author	Thomas W. Goolsby
Year	1994a
Purpose	• To investigate selected aspects of perceptual processing of successful musicians; to determine the extent that eye movements of music readers reflect perceptual processing • To determine the extent to which eye movement differs between skilled music readers and less-skilled music readers
Measurement	• Measuring the relative music-reading ability of subjects at two levels, the complexity of notation, the encounter with music in three ways Dependent variables: number of progressive fixations, progressive fixation duration, progressive saccade length, number of regressive fixations, regressive fixation duration, regressive saccade length
Assessments Used	Pretest: Belwin-Mills Singing Achievement Test (Bowles, 1971) Questionnaire with 24 demographic variables
Description	
Subjects	24 graduate students in school of music selected; 12 with highest scores and 12 with lowest scores from 72 volunteers, after auditioning with singing test
Instruments Used	Voice
Stimuli	Four melodies had various levels of notational complexity Melodies selected from *Solfege des Solfeges* (Danhauser, Lemoine, & Livigna, 1910): single-line, four staves, treble clef, C major, 4/4 meter. The physical dimensions of the melodies were held constant, with varying amounts of visual detail within the space. The complexity of the melodies was evaluated by 12 music faculty: Level A – 51 half and whole notes, 3 half rests, 4 dynamic markings, 3 breath marks Level B – 76 quarter and half notes, 6 quarter and half rests, 7 tie/slur markings, 1 dynamic marking, 1 breath mark, 6 accents Level C – 83 eighth, quarter, half, whole notes, 1 quarter rest, 18 tie/slur markings, 8 breath marks, 1 fermata, 3 tempo changes Level D – 65 eighth, quarter, half, whole notes, 5 half rests, 13 accidentals, 3, ties, 11 dynamic markings, 10 breath marks, 2 expression markings
Apparatus	Eye movement measured with Stanford Research Institute *Dual Purkinje Image Eyetracker* (SRI) Computer used to sample horizontal and vertical components of eye position every millisecond, and the location of each fixation Slide projector presented visual stimuli (melodies) Cassette recorder for performances Metronome for tempo Partial bite plate for head position stability
Method	The music students read the four melodies with an initial performance (sight-reading), an immediate second encounter, and a final performance after a short period of free study/practice Each melody was performed three times while eye movements were monitored and recorded Subjects could vocalize melodies in a comfortable manner (e.g. solfege, humming, “ah”ing) while maintaining tempo
Results	There was an eye-movement difference found between groups with progressive fixation duration Skilled readers used more but shorter fixations Mean saccade lengths were longer for less-skilled readers Suggests that skilled reader use more eye movement to look ahead, and then regress more to return to the point of performance Spacing of notation affects eye movement regardless of performance accuracy

Reliability	-
Validity	Belwin-Mills Singing Achievement Test publishes a concurrent validity correlations of .81 with Aliferis Music Achievement Test and .73 with WFPS
Reference	Goolsby, T. W. (1994a). Eye movement in music reading: Effects of reading ability, notational complexity, and encounters. *Music Perception, 12*(1), 77-96.

Title	Profiles of processing: Eye movements during sight reading
Author	Thomas W. Goolsby
Year	1994b
Purpose	• To investigate the extent of differences in music perception, or the processing of musical notation, by skilled and less-skilled sight readers • To find insight into cognitive processing of musical notation with regard to control of eye movement profiles
Measurement	• Performance errors, expression, musicality • Number of progressive and regressive fixations, progressive and regressive fixation duration, progressive and regressive saccade length The recordings were measured by 3 music professors
Assessments Used	-
Description	
Subjects	2 subjects, representing two distinct levels of sight reading ability from an initial experiment with 24 subjects
Instruments Used	Voice
Stimuli	3 melodies, 4, staves, treble clef, C major, 4/4 meter. They varied in notational complexity, but had the same physical dimensions. Level A: 54 notes/rests, 4 dynamic markings, 3 breath marks. Level B: N/A Level C: 84 notes/rests, 18 tie/slur markings, 8 breath marks, 3 tempo changes. Level D: 70 notes/rests, 13 accidentals, 3, ties, 13 dynamic markings, 10 breath marks
Apparatus	Stanford Research Institute *Dual Purkinje Image Eyetracker* (SRI) for monitoring eye movement Movement recorded by *Digital Equipment Corporation PDP-11/40* computer Slide projector to present the melodies Microphone and cassette record used to record performances
Method	After the tempo was provided, each subject vocalized each melody as they were comfortable
Results	Less-skilled reader: used more progressive fixations, but slightly fewer regressions. The less-skilled reader spent more time fixating at the beginning of each melody, and had many consecutive horizontal fixations. Less-skilled reader fixated on almost every note and most rests. Skilled reader: longer progressive saccade lengths. The skilled reader looked farther ahead, and then back to the point of performance. Skilled reader was more accurate and more musical. He had more saccades between staves, and did not fixate on every note, nor did he fixate near the dynamic markings, even though they were performed. Suggests that skilled readers perceive visual information through peripheral vision.
Reliability	-
Validity	-
Reference	Goolsby, T. W. (1994b). Profiles of processing: Eye movements during sight reading. *Music Perception, 12*(1), 97-123.

Title	The effect of tonal pattern training on the aural perception, reading recognition, and melodic sight-reading achievement of first-year instrumental music students
Author	Patricia Ann Grutzmacher
Year	1987
Purpose	• To investigate the relationship of tonal pattern instruction using harmonization and vocalization to tonal concept development and the performance achievement of beginning wind instrumentalists • To compare two courses of study: one emphasizing tonal concept development and the other with a single-note identification approach
Measurement	• Sight-reading achievement • Aural perception of tonal patterns in major and minor modes • Reading recognition of major and minor tonal patterns perceived aurally and visually, simultaneously
Assessments Used	Pretest: MAP Tonal Imagery Pretest and posttest: Iowa Tests of Musical Literacy (ITML) Level 2, Tonal Aural Perception (Gordon, 1970) Iowa Tests of Musical Literacy (ITML) Level 2, Tonal Reading Recognition Posttest: Author-composed Melodic Sight-reading Achievement Test (MSRAT)
Description	
Subjects	48 grade 5 and 6 brass or woodwind students in 3 elementary schools. Randomly assigned to instrumental classes, and each class randomly assigned as experimental or control group
Instruments Used	Voice (for training), brass and woodwind Instruments
Stimuli	Experimental Group Content: 10 major-key and 10 minor-key tonal patterns, as well as Alfred's Basic Band Method (1977) Control Group Content: Same text used but no tonal patterns
Apparatus	-
Method	Lessons included a 30-minute class per week for 14 weeks. All taught by the researcher. Experimental Group: 10-minute warm-up where tonal material presented aurally and then notationally. The activities were long tones, scales, arpeggios using harmonization and vocalization with syllables. Tonal patterns were isolated and vocalized. Then, recognition of familiar tonal patterns and presentation of new fingerings for these tonal patterns. Finally, playing activities at aural level and music reading. Control Group: 10-minute warm-up with only notational presentation of tonal material. All material presented with definitions and descriptors. Each exercise in text performed without vocalization. New material was presented by notation, fingering, and playing concurrently
Results	Found a significant difference in the posttest melodic sight-reading achievement mean scores, with the experimental group scoring higher. The experimental group scored higher on the ITML test, showing they were better able to aurally identify major and minor tonalities. Both groups improved with aural and visual comparison of tonal patterns, but not shown to be due to treatment. Experimental group showed significant improvement from MAP to posttest ITML. Found that tonal pattern instruction with vocalization improved melodic sight-reading skills. Teaching major and minor through experience (over only definition) improves conceptual understanding. There needs to be a balance of aural-visual perception skills
Reliability	-
Validity	-
Reference	Grutzmacher, P. A. (1987). The effect of tonal pattern training on the aural perception, reading recognition, and melodic sight-reading achievement of first-year instrumental music students. *Journal of Research in Music Education, 35*(5), 171-181.

Title	Pitch error analysis of young piano students' music reading performances
Author	Helga Rut Gudmundsdottir
Year	2010a
Purpose	• To explore pitch reading errors of young piano students and the effect of age of the error making • To propose error categories for use as basis for further studies on music reading
Measurement	• Pitch Error Types: Erroneous pitches, redundant pitches, and omitted pitches • Errors found in RH or LH • Self-corrections
Assessments Used	-
Description	
Subjects	35 pianists ages 6-13, in second or third year of piano study. Younger group (n=18, age range=6-9 years). Older group (n=17, age range=10-13 years)
Instruments Used	Piano
Stimuli	Three author-composed pieces, in C major, 8 measures in length. Melodies moved by seconds and thirds and rhythms had only eighth and quarter notes. Piece A: melody alternated between RH and LH. Pieces B and C: identical with same number of seconds and thirds, with different LH complexity
Apparatus	Digital piano for performances Sequencing software on a computer for recording performances
Method	Subjects received 30-minute lesson once a week, and played through the three pieces in random order
Results	Found that pitch reading skills develop through childhood regardless of formal training. In both groups, most common error was erroneous pitches, then redundant pitches. Omitted pitches were few. Primary goal for majority of subjects was to play pitches correctly at expense of timing accuracy. Younger group make more incorrect pitches and a higher number of redundant pitches than older group. Younger group made more redundant errors where LH and RH played simultaneously. With both groups, more errors occurred in the LH
Reliability	-
Validity	-
Reference	Gudmundsdottir, H. R. (2010a). Pitch error analysis of young piano students' music reading performances. *International Journal of Music Education, 28*(1), 61-70.

Title	Let's fact the music: A behavioral and electrophysiological exploration of score reading
Author	Thomas C. Gunter, Bjorn-Helmer Schmidt, and Mireille Besson
Year	2003
Purpose	• To investigate the cognitive aspects underlying music reading by means of electrophysiology • To explore whether sight reading of diatonic violations in a musical phrase elicits EEG reflections similar to those when hearing diatonic violations
Experiment 1: Self-paced reading of musical score	
Measurement	• Behavioral: Reading time of notation with and without diatonic violations
Assessments Used	-
Description	
Subjects	8 musicians from advanced music school with 17-26 years of classical Western music training
Instruments Used	-
Stimuli	120 familiar music phrases selected from materials of Besson and Faita (1995) study. Every phrase had a version with a diatonic violation in middle of phrase. The diatonic violation was between 2 and 12 measures after beginning of the score. Half of phrases contained violation. The musical phrase was presented first with clef, key, and time signatures. This information stayed until the phrase ended

Apparatus	Color monitor for displaying scores Button to change visual stimuli
Method	There was a short training session with 5 musical phrases. Then, for one hour, the subject saw one measure at a time, and pressed a button to move to the next measure. When the end of the score was reached, the subject had to indicate whether or not a violation was present in the score and if they knew the piece from which the score was derived
Results	Self-paced reading times were between 1700 and 300 ms. With measures containing violations, it was clear that a significant difference in reading time was found – around 500 ms slower than correct measures. As the sequence unfolded, subjects' reading times became faster overall
Reliability	-
Validity	-
Experiment 2: ERPS during reading of musical score	
Measurement	• Bipolar horizontal EOG, bipolar vertical EOG
Assessments Used	-
Description	
Subjects	11 RH musicians (mean age 22.5 years) at advanced music school with 13-25 years of training in classical Western music
Instruments Used	-
Stimuli	Same materials as used in experiment 1
Apparatus	Color monitor for displaying scores, using an RSVP procedure (3000 ms by measure and 200 ms blank screen in between) Button for responding Electroencephalogram (EEG) and electrodes on electro-cap for recording
Method	3-hour sessions, with subjects in a dimly lit room, facing the color screen. They were to read the score and answer 2 post-score questions: Was there a violation? Do you know this piece of music? They pressed a button for yes/no for each answer. They were told to blink after the post-score tasks
Results	They found that with the RSVP format, the ERPs for the measures with violations showed an N1 component (this reflects modality-specific processing in the visual field) followed by a long-lasting negative shift that encompassed three effect: 1. Early Negativity, 2. Middle Latency Negativity, 3. Long Lasting Negativity Violation-detection in music relies on modality-independent and abstract processing mechanisms. In early negativity, the detection is comparable to what is found in auditory music processing. In later time frames, the difference between sight reading and hearing music are apparent in the ERPs
Reliability	-
Validity	-
Reference	Gunter, T. C., Schmidt, B. H., & Besson, M. (2003). Let's face the music: A behavioral and eletrophysiological exploration of score reading. *Psychophysiology, 40*(5), 742-751.

Title	Instrumental musical performance: One approach toward evaluation
Author	Kenneth U. Gutsch
Year	1966
Purpose	• To determine if an evaluative tool could be developed which would measure an individual's instrumental music achievement while sight-reading rhythms • To determine whether or not such a tool could differentiate degrees of attainment for individuals who represented different amounts of instrumental music experience and reflected a variety of age levels
Measurement	Administration and scoring of test is not described in detail
Assessments Used	-
Description	
Subjects	771 students grade five or older, with at least 6 weeks instrumental music experience
Instruments Used	-
Stimuli	200 flash cards, 100 of each form (Form A or Form B) of rhythmic problems, each consisting of one or more measures in length
Apparatus	Tape recorder for performances
Method	The test was administered individually, with 100 flash cards presented to a subject, and then the other 100 flash cards, within a 48 hour period One year later, the test was re-administered to 81 of the previous subjects The cards were reordered and administered to 120 students who had not previously seen the test
Results	The test reflected test reliability and scorer reliability Experience was the most influential factor governing performances Age had little effect upon sight-reading rhythm ability
Reliability	Scorer reliability .99
Validity	Results deem that the test has construct validity
Reference	Gutsch, K. U. (1966). Instrumental music performance: One approach toward evaluation. *The Journal of Educational Research, 59*(8), 377-380.

H

Title	Musical expertise and melodic structure in memory for musical notation
Author	Andrea R. Halpern, and Gordon H. Bower
Year	1982
Purpose	• To investigate the role of melodic structure on short-term memory for musical notation by musicians and non-musicians
Experiment 1: Pilot	
Measurement	Unknown
Assessments Used	-
Description	
Subjects	24 subjects, 12 of which were musicians and 12 of which were non-musicians
Instruments Used	-
Stimuli	Used good and bad melodies, 10 quarter notes each, without accidentals. 6 graduate students classified the melodies as "better" or "poorer". 36 melodies in total. There were also 18 'random' melodies composed. To determine whether the melodies were "good" or "bad", 13 musicians and 9 non-musicians listened to a tape recording of the melodies played on an organ. They rated each melody on a 9-point scale
Apparatus	Cards for melodies Slides for presenting the cards Projector for viewing slides Paper and pencil for recall
Method	Subjects saw a slide with a melody for 5 s, and recalled it in written form 15 s after the slide had been removed. This 15 s interval either had a distracting auditory or visual task
Results	Musicians performed better than non-musicians Good melodies were remembered better than bad melodies Melody type interacted with subject type
Reliability	-
Validity	-
Experiment 2	
Purpose	• To investigate the effect of musical structure on remembering notation
Measurement	• Each melody was scored on an absolute basis, where every correct note was scored, but one incorrect note resulted in an overall score of 0 • The recall patterns were also scored by a similarity judgment procedure, where 'two musically knowledgeable' judges compared the responses of 3 musicians and 3 non-musicians to the correct answer, assigning it a similarity score from 1 (low) to 10 (high). They used any criteria of visual/musical similarity that they felt was relevant. Both systems revealed the same pattern of results
Assessments Used	During experiment, questionnaire for musical background
Description	
Subjects	12 subjects, with 6 student musicians with at least 10 years of experience with reading music, and 6 students with no music reading experience
Instruments Used	-
Stimuli	The good, bad, and random melodies were used form the pilot experiment. They appeared on a staff with a treble clef, and all consisted of 10 quarter notes
Apparatus	Index cards to present melodies Photographic slides of each card for viewing Staff paper for transcription Pencil
Method	1-3 subjects were tested at a time. They began with 6 practice trials. The experimenter explained the experiment, said 'ready', and activated a timer. The stimulus slide was projected for 4 s, and removed. The subjects wrote down the melody on staff paper, and were encouraged to recall 10 notes. They had unlimited time for recall. They saw

	18 good, 18 bad, and 18 random melodies in a random order
Results	Musicians recalled good melodies better than bad melodies. This shows that the influence of melodic structure occurs early in the music reading process. Non-musicians recalled good and bad melodies at the same moderate level. Both groups recalled random melodies more poorly
Reliability	Inter-judge reliability = .97
Validity	-
Experiment 3	
Purpose	• To assess the importance that ease of grouping a melody into sub-patterns has for music reading • To investigate how many groupings melodies are divided into
Measurement	• The efficiency and number of sub-patterns in a melody • Consistency with other subjects as to their melodic segments. Consistency was established if at least 6 of the subjects agreed about the presence or absence of a division • The number of groupings for each melody
Assessments Used	-
Description	
Subjects	14 subjects, 7 of which were students and professional musicians, all fluent in music reading, and 7 who were students with no knowledge of music reading
Instruments Used	-
Stimuli	The good and bad melodies from the previous experiments, with 10 notes each. A total of 36 melodies were used
Apparatus	Sheet of paper for presentation of stimuli Writing utensil
Method	The subjects saw 18 good melodies and 18 bad melodies on a sheet of paper in random order. They were instructed to divide the notes into groupings by marking the melody with zero, one, or more vertical slashes. They were told to use any criteria for the division, and they could have as many divisions as they liked
Results	The musicians did not group good melodies more consistently than the bad melodies. They tended to group both melodies idiosyncratically Musicians saw fewer groupings in a good melody than a bad melody
Reliability	-
Validity	-
Reference	Halpern, A., & Bower, G. (1982). Musical expertise and melodic structure in memory for musical notation. *American Journal of Psychology, 95*, 31-50.

Title	Effect of delayed auditory feedback on musical performance
Author	Larry L. Havlicek
Year	1968
Purpose	• To investigate the effect of auditory feedback, with abnormally delayed auditory feedback, on musical performance • To determine whether or not the same effect would occur during music performance
Measurement	• Number of performance errors, defined as stopping or hesitating, playing a wrong rhythmic passage, playing wrong notes, adding or leaving out passages of music. When there was a series of errors, only one instance of error counted for the series • Reaction to synchronous and delayed auditory feedback
Assessments Used	-
Description	
Subjects	21 university students in music education
Instruments Used	Piano, brass, woodwind, string
Stimuli	2 pairs of melodic etudes from Gariboldi's *Twenty Studies for Flute*, arranged for each instrument used For pianists, 4 short excerpts from comparable selections chosen from *The Etude* music magazine of 1936
Apparatus	*Magnecord Model M-30* magnetic tape record for delayed auditory feedback *Permoflux Dynamic* earphones Microphone for recording
Method	Each subject performed 4 short compositions in a sound-treated laboratory while listening to the sound of their instrument through earphones. The subject sat and put on earphones. The subject played a few notes on their instrument to begin, while the recording and headphones were adjusted for the experiment. The subject sight-read the 4 compositions, with the first 2 with synchronous feedback, and the second 2 with delayed auditory feedback. The subject's remarks about the adjustment to delayed feedback was also recorded after their performances
Results	During music performance under delayed auditory feedback, there are significantly more errors The effect of delayed auditory feedback are equally disturbing for all classifications of musical performance used in this experiment The effects of delayed performance are similar to the effects on speech, including a general slowing down, an increased number of articulatory errors, occasional stops or blocking, and occasional additions of extraneous material
Reliability	-
Validity	-
Reference	Havlicek, L. L. (1968). Effect of delayed auditory feedback on musical performance. *Journal of Research in Music Education, 16*(4), 308-318.

Title	Relationships among music sight-reading and technical proficiency, spatial visualization, and aural discrimination
Author	Carol M. Hayward and Joyce Eastlund Gromko
Year	2009
Purpose	• To examine the strength of predictors of music sight-reading ability; technical proficiency (kinesthetic/aural), aural discrimination of tonal and rhythmic patterns (aural/spatial), spatial-temporal reasoning (visual/spatial)
Measurement	• Music sight-reading ability (with WFPS, Form A). This was measured by one graduate student, for pitch and rhythm errors • Technical proficiency (with author-composed test) • Aural discrimination of patterns (with AMMA) • Spatial visualization (with 3 subtests from Kit of Factor-Referenced Cognitive Tests)
Assessments Used	WFPS, Form A (Watkins & Farnum, 1954) Author-composed adaptation of test of Elliott (1982), with ascending and descending 1-ocatave major and minor scales, ascending and descending 1-octave arpeggios, chromatic scales AMMA (Gordon, 1989) 3 subtests from Kit of Factor-Referenced Cognitive Tests (Ekstrom, French, Harman, & Dermen, 1976)
Description	
Subjects	70 university students, 62 of whom were music majors
Instruments Used	Woodwind instruments, brass instruments (according to the subject's instrument)
Stimuli	(Use of Assessment Materials)
Apparatus	Paper Pencil
Method	Subjects were tested individually with the WFPS, Form A, for sight-reading ability, followed by an author-composed test for technical proficiency. They were tested in groups with AMMA for aural discrimination of patterns and with 3 subtests of the Kit of Factor-Reinforced Cognitive Tests
Results	Found that, true to their hypothesis, speed and accuracy of music sight-reading could be predicted by a combination of aural pattern discrimination, spatial-temporal reasoning, and technical proficiency. Aural discrimination and spatial-temporal reasoning loaded together on one factor separate from technical proficiency, which suggests that sight-reading involves reading and playing
Reliability	WFPS = .87-.94 AMMA = .85 (split halves), .87 (retest) 3 subtests of Kit of Factor-Referenced Cognitive Tests = .81, .84, .90
Validity	-
Reference	Hayward, C. M., & Eastlund Gromko, J. (2009). Relationships among music sight-reading and technical proficiency, spatial visualization, and aural discrimination. *Journal of Research in Music Education, 57*(1), 26-36.

Title	The development of a vocal sight-reading inventory
Author	Michele Henry
Year	2001
Purpose	• To develop an individualized vocal sight-reading assessment tool, to represent skills involved in vocal sight reading. Also, a tool that will accommodate time constraint in high school choral setting, allows formative and summative evaluation, and provides validity, reliability, and normative data. Main sub-questions: A. What pitch reading skills should be assessed? B. How should these skills be assessed? C. What should be the testing and scoring procedure? D. Will these procedures be efficient enough to merit use in the choral curriculum of secondary schools? E. How is test reliability to be established?
Measurement	• Only pitch accuracy was measured • VSRI intended for individual assessment of secondary-level singers Scoring: Evaluated only pitches within identified component skill patterns (remaining pitches were not scored). Scoring sheets used numeric codes, attached to pitches. Scorers identified incorrect pitches by placing mark over the numeric code of incorrect pitch. Subject received score of number of correct component skills (0-28). Each performance also scored note-by-note, with credit for each correct pitch. Subject received score of number of correct notes (0-137 Form A; 0-151 Form B). All performance tapes reviewed independently by researcher and another scorer (Note: researcher's scores used for statistics, and second scores used only for inter-scorer reliability) For Pilot Study: A second scorer rated with component skill method.
Assessments Used	The Vocal Sight-Reading Inventory (VSRI) was created and tested
Description	
Subjects	Pilot study of test: 42 high school students Main test: 183 students from 5 high schools (10 students from each school were randomly tested)
Instruments Used	Voice (Piano for experimenter)
Stimuli	Tonal melodies created to represent the level of secondary choral music Test items consisted of entire melodies in major keys, with embedded pitch skills Pitch skills assessed were derived from patterns found in tonal music and identified by their function within a harmonic framework 7 skill categories established for VSRI: Melodic patterns representing single tonal harmonic or scalar function, including A. Conjunct (includes repeated notes and scale patterns spanning third, fifth, seventh), B. Tonic, C. Dominant, D. Subdominant, E. Cadential, (4 authentic patterns) F. Modulatory, G. Chromatic Conjunct patterns further classified by direction of ascending or descending motion. Triadic 2-note patterns implied specific harmonic function and further categorized by general interval size (labeled 'skips' or 'leaps'). 2 forms of the test: Each contained 8 melodies, with each melody containing several pitch skills. Most melodies were I-IV-V-I or I-V-I-IV-V-I harmonic progression. Melodies in keys of C, F, and D – written in both treble and bass clefs. All melodies were quarter, half, dotted half notes – therefore no complicated rhythms.
Apparatus	-
Method	Creation of VSRI followed this criteria: A. Criterion-based rather than norm-based, B. Assess skill in pitch reading only, C. Its items would consist of entire melodies in major keys, D. Embedded within melodies would be tonal patterns representative of discrete sight-singing skills, E. Scoring procedures be efficient and reliable. Scores from the 'new' system were compared with those obtained by note-by-note scoring. During Test: Subject, in practice room, was given instructions, and given melodies. For each melody, key was established and first pitch played. Subject had 30 seconds to study/practice melody. Then chord progression and first note repeated and subject

	sang melody
Results	Subjects scored poorly on VSRI sight-singing test, though this could be because of large number of subjects, or for other reasons unknown to the researcher. Found that scoring by identified pitch skills was adequate to represent the skills required for sight-reading tonal music
Reliability	Inter-scorer reliability r = .97 Correlation between the 2 compared scoring systems r = .96
Validity	-
Reference	Henry, M. (2001). The development of a vocal sight-reading inventory. *Bulletin of the Council for Research in Music Education, 150*, 21-35.

Title	The use of targeted pitch skills for sight-singing instruction in the choral rehearsal
Author	Michele L. Henry
Year	2004
Purpose	• To determine the effect of emphasizing specific pitch skills based on scale degree and harmonic function to teach sight-singing in the choral rehearsal
Measurement	• Local sight reading ability • Pretest: Independently scored by the researcher and a graduate student assistant • For VSRI, every note in each melody was assessed. For skill patterns, each received one point, with a possible total score of 15
Assessments Used	15 pitch skills were identified, including repeated pitches, ascending and descending scalar motion, skips and leaps within the tonic, dominant, and subdominant triads, cadential patterns Pretest: portion of the Vocal Sight Reading Inventory (VSRI) (Henry, 2001) that incorporated these pitch skills Posttest: full VSRI administered
Description	
Subjects	67 subjects, divided into two intact groups: Group A: Solfege Drill (n=41), Group B: Familiar melodies (n=26). All were high school choral students
Instruments Used	Voice
Stimuli	Group A: Teacher-constructed exercises with the targeted pitch skills. All melodies for written with solfege syllables (no formal music notation) placed vertically according to the relative ‘height’ of the pitch Group B: 12 familiar melodies, as determined by an informal survey by the conductor. Each pitch skill was represented at least once in the melodies. Melodies had standard notation with solfege syllables next to each note head
Apparatus	Digital minidiscs for pretest recording
Method	During a 12-week period, both groups received 3 treatment sessions per week. Each session was 5-8 minutes. The same teacher taught each group. Group A: For the first 6 weeks, they were trained with the 15 targeted pitch skills using solfege drills (without hand signs). For the last six weeks, one skilled category was the ‘target skill for the day’. The skill was introduced by solfege syllables. Group B: For first 6 weeks, students reviewed the songs with text, but then with only solfege syllables (without hand signs) until memorized. For the last 6 weeks, one skill category was the ‘target skill for the day’. The skill was introduced by solfege syllables
Results	Subjects with instruction with targeted pitch skills achieved a significant increase in their sight-singing ability. Group A: Had significant gains in ability. Group B: Had significant gains in ability
Reliability	Inter-score reliability = .97
Validity	-
Reference	Henry, M. L. (2004). They use of targeted pitch skills for sight singing instruction in The choral rehearsal. *Journal of Research in Music Education, 52*(3), 206– 217.

K

Title	Saccadic eye movements while reading music
Author	Veronica Kinsler, R. H. S. Carpenter
Year	2000
Purpose	• To use music as a stimulus for probing the mechanisms of oculomotor scanning
Measurement	• Subjects' eye movements
Assessments Used	-
Description	
Subjects	4 subjects; all competent musicians. 2 were well-informed to purposes of experiment and 2 were not
Instruments Used	- (Tapping only)
Stimuli	Lines of simplified music with rhythmic information only, in conventional musical notation. Single line of notes and bar lines. Stimulus display was green.
Apparatus	Computer screen for presentation Second computer for recording output Microphone for response Scleral infra-red eye movement transducer (Carpenter, 1988) with bitebar *Epic* physiological recording system (Cambridge Research Systems, Ltd.) for overall control of experiment, including triggering of stimuli and recording of signals from the eye-movement transducer and microphone (this program ran on second computer) Metronome to signal the tempo
Method	Single overall trial with 32 trials; each was 10 s with inter-trial gap of 5 s. Subjects randomly presented with a stimulus on computer screen. Subjects tapped the rhythmic phrase on a microphone. Before and between trials, they fixated on a spot.
Results	Found that there is tendency to fixate on significant details of notation (e.g. notes and bar lines) rather than spaces in between. Shorter notes less likely to be fixated than longer ones. Found that the time of execution of individual saccades seems to be unrelated to time of execution of elements of the performance itself. As the tempo is increased, the average time between saccades decreases, but their mean amplitude increases. Authors propose a model of oculomotor and perceptual processes, with a fixated image internally scanned and interpreted to a given criterion of accuracy. The scan ends with the criterion cannot be reached, and this determines the position of the next fixation.
Reliability	-
Validity	-
Reference	Kinsler, V., & Carpenter, R. H. S. (1995). Saccadic eye movements while reading music. *Vision Research, 35*(10), 1447-1458.

Title	The musicians' glance: A pilot study comparing eye movement parameters in musicians and non-musicians
Author	Reinhard Kopiez and Niels Galley
Year	2002
Purpose	• To test whether the eye movements (saccades) of professional musicians differ from a corresponding non-musician control group
Measurement	• Saccades • Tapping • Mental Speed (with number connection test) • EOG tasks for recording the changing electrical field with eye movement (sinusoidal movement, rectangular movement, counter movements), Baseline EOG (with interview), EOG with closed eyes (summarized as psychomotoric optimization strategies, mental speed, general mental capacity)
Assessments Used	Tasks: 1. Tapping of R and L hands, 2. Number Connection Test (Oswald & Roth, 1997), 3. Looking at picture, 4. Horizontally running point, 5. Horizontally jumping point (1), 6. Horizontally jumping point (2), 7. Vertically jumping point, 8. Horizontally jumping point (3), 9. Horizontally counter-jumping point, 10-21. Raven-SPM, Series D (Raven, 2000), 22. Interview, 23. Picture remembering
Description	
Subjects	Experimental Group: 8 professional musicians from a music academy. Control Group: 254 psychology students
Instruments Used	-
Stimuli	Oculomotoric tracking task with a horizontal shifting dot ("jumping point"), whose velocity increased over 90 seconds from 0.2-1.5 Hz, following a rectangular waveform of movement
Apparatus	*PAR electronic*, Berlin EOG amplifier device for recording the changing electrical field caused by the moving dipole of the eye Skin electrodes to subjects' face on horizontal and vertical axis Computer monitor to show the jumping point Head rest for subject's head
Method	Subjects sat in front of monitor and participated in a sequence of 23 tasks, as described above in "assessments used"
Results	Only response saccades from the horizontally jumping point were analyzed. Data showed clear differences in eye movement parameters between musicians and non-musician. The 'musicians' glance' was characterized by: considerably reduced frequency of misses, shorter reaction times in reactive saccades, high proportion of anticipatory saccades, higher saccadic velocity, shorter anticipatory latencies. Professional musicians had efficient strategies for processing of an oculomotoric task
Reliability	-
Validity	-
Reference	Kopiez, R., & Galley, N. (2002). The musicians' glance: A pilot study comparing eye movement parameters in musicians and non-musicians. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, & J. Renwick (Eds.), *Proceedings of the 7th International Conference on Music Perception and Cognition* (pp. 683-686).

Title	Towards a dynamic model of skills involved in sight reading music
Author	Reinhard Kopiez and Ji In Lee
Year	2006
Purpose	• To reveal relationship between complexity of SR stimuli and predictor skills in order to build up a level-specific 'dynamic' model of SR in addition to the general model. • To investigate the relationship between selected predictors of achievement in sight reading and changing complexity of sight reading tasks. How do different variables gain or lose significance as sight reading stimuli becomes more difficult?
Measurement	Measurement of 23 selected skills in 3 categories: • General cognitive skills: Working memory, Short-term memory (STM), Short-term music-specific memory (STMM), General mental capacity (with Raven D) • Elementary cognitive skills: Speed of information processing (NCT), Simple visual reaction time (RTV), Simple auditory reaction time (RTA), Tapping speed, Trilling speed • Expertise-related skills: Accumulated hours of solo practice up to the age of 10, 15, 17, and total, Accumulated hours of piano lessons up to the age of 10, 15, 17, total, Accumulated hours of sight-reading expertise up to the age of 10, 15, 18, total, Inner hearing Scoring of SR done by researcher-developed software program, *Midicompare*, which matches performance to score with adjustable window time. The software calculates the number of correct notes (matched score notes), missed notes and extra notes. In this study, however, only matched score notes (matches) are used for achievement analysis.
Assessments Used	Raven's D Matrix (Raven, 2000)
Description	
Subjects	52 piano major graduates and undergraduates from Hanover University of Music and drama
Instruments Used	Piano
Stimuli	For sight reading component: 5 pieces of increasing complexity and of similar length, selected from the University of South Africa exam syllabus for piano sight reading (UNISA, 1995). They were arranged for solo voice and bi-manual piano accompaniment by a professional composer. There was pre-recorded pacing with the solo voice performed by a violinist playing with a metronome. Pieces were similar length and arranged for solo voice and 2-handed piano accompaniment by professional composer. Task complexity of pieces was evaluated by external judges.
Apparatus	Metronome for introduction to pieces *MIDI* keyboard Loudspeaker for presentation of solo track *Midicompare* (Dixon, 2000) for scoring of SR
Method	Two warm-up pieces to start, then 5 pieces of increasing complexity. Subjects given 60 s to review piece before performing. They heard a metronome click to signal the solo melody beginning. Working Memory: subjects added/subtracted in steps of one in a 3x3 matrix, 18 test exercises Short-term Memory: numerical test where numbers were shown one at a time and subjects had to remember Short-term Music-specific Memory: subjects looked at 12-bar melody for 1 minute and then try to play as many correct notes as they could from memory Raven's D Matrix: 12 items of series D used Number Combination Test: subjects joined the numbers 1-90 in ascending order as quickly as possible Auditory and Visual Reaction Time: subjects had to release a morse key as soon as they saw or heard the stimuli. 20 test exercises

	Speed Trilling: using different finger combinations, subjects had to trill for 15 seconds. Both trills were repeated once Speed Tapping: subjects tapped for 30 s on a morse key Inner Hearing: used embedded melody paradigm (Brodsky et al., 2003). Subjects saw the variation of a theme for 45 s. They had to imagine the sound without humming or singing the score. They then heard the theme OR the lure melody through the speakers. They had to decide whether or not the theme was embedded in the variation seen Retrospective Interviews: subjects described the number of hours for solo practice, sight reading, years of piano lessons
Results	With increasing task demands, sight reading ability is determined by both practice dependent skills and skills which are assumed to be limited by innate abilities such as psychomotor movement speed. Sight reading achievement is explained as result of specific combinations of different categories of skills which change with the demands of a task. Specifically, a combination of speed of information processing, psychomotor speed and SR expertise.
Reliability	-
Validity	-
Reference	Kopiez, R. & Lee, J. I. (2006). Towards a dynamic model of skills involved in sight reading music. *Music Education Research, 8*(1), 97-120.
See Also	Kopez, R., Weihs, C., Ligges, U., & Lee, J. I. (2006). Classification of high and low achievers in a music sight-reading task. *Psychology of Music, 34*(1), 5-26.
See Also	Kopiez, R., & Lee, J. I. (2008). Towards a general model of skills involved in sight reading music. *Music Education Research, 10*(1), 41-62.
See Also	Lee, J. I. (2004). *Component skills involved in sight reading music*. Frankfurt am Main: Peter Lang.

Title	Towards a general model of skills involved in sight reading music
Author	Reinhard Kopiez and Ji In Lee
Year	2008
Purpose	• To develop a general model of sight reading which can explain observed achievement variance with a limited number of predictor variables covering many levels of information processing • To give insight into the underlying predictor structure by identifying basic components and factors that group selected predictors
Measurement	Measurement of 23 selected skills in 3 categories: • General cognitive skills: Working memory, Short-term memory (STM), Short-term music-specific memory (STMM), General mental capacity (with Raven D) • Elementary cognitive skills: Speed of information processing (NCT), Simple visual reaction time (RTV), Simple auditory reaction time (RTA), Tapping speed, Trilling speed • Expertise-related skills: Accumulated hours of solo practice up to the age of 10, 15, 17, and total, Accumulated hours of piano lessons up to the age of 10, 15, 17, total, Accumulated hours of sight-reading expertise up to the age of 10, 15, 18, total, Inner hearing Scoring: Evaluation method that is a. objective, b. repeatable and independent of the raters' reliability, c. based on an automized procedure. To fulfill this, they used a software-based solution that counts number of correctly performed pitches within a given timeframe. Only percentage of total score matches for both hands were considered
Assessments Used	Raven's D Matrix (Raven, 2000)
Description	
Subjects	52 piano major university students or graduates, including undergraduates, postgraduates, professional accompanists, winners of sight-reading competitions

	(mean age 24.56 years). 24 males and 28 females
Instruments Used	Piano
Stimuli	For sight reading component: 5 pieces of increasing complexity and of similar length, selected from the University of South Africa exam syllabus for piano sight reading (UNISA, 1995). They were arranged for solo voice and bi-manual piano accompaniment by a professional composer. There was pre-recorded pacing with the solo voice performed by a violinist playing with a metronome
Apparatus	Loudspeaker to play the solo voice *MIDI* keyboard for performances *CUBASE 5.1* (Steinberg, 2000) for sequencing program Recorder for performances *Midicompare* (Dixon, 2002) for data analysis of sight-reading achievement *PowerPoint* for presentation Morse Key for various tests
Method	Working Memory: subjects added/subtracted in steps of one in a 3x3 matrix, 18 test exercises Short-term Memory: numerical test where numbers were shown one at a time and subjects had to remember Short-term Music-specific Memory: subjects looked at 12-bar melody for 1 minute and then try to play as many correct notes as they could from memory Raven's D Matrix: 12 items of series D used Number Combination Test: subjects joined the numbers 1-90 in ascending order as quickly as possible Auditory and Visual Reaction Time: subjects had to release a morse key as soon as they saw or heard the stimuli. 20 test exercises Speed Trilling: using different finger combinations, subjects had to trill for 15 seconds. Both trills were repeated once Speed Tapping: subjects tapped for 30 s on a morse key Inner Hearing: used embedded melody paradigm (Brodsky et al., 2003). Subjects saw the variation of a theme for 45 s. They had to imagine the sound without humming or singing the score. They then heard the theme OR the lure melody through the speakers. They had to decide whether or not the theme was embedded in the variation seen Retrospective Interviews: subjects described the number of hours for solo practice, sight reading, years of piano lessons
Results	General cognitive skills: No predictors in this group were significant. Sight reading does not seem to be influenced by memory capacity of aspects of intelligence Elementary cognitive skills: Trilling speed (especially with finger combination 3-4) and speed of information processing were predictors for sight reading achievement. Trilling speed represents the intersection between task-specific training and practice-independent advantages in movement speed. Speed of information processing is the only predictor assumed to be independent of practice Expertise and practice-related skills: There is crucial time window for acquisition of sight-reading expertise. Number of accumulated hours of sight-reading practice up to age of 15 is best predictor for sight-reading achievement General: Sight-reading expertise is determined by the time spent on activities related to this skill. There is a critical time window before turning 15 for optimal training. Fluency of sight reading is more influenced by mental speed than by memory capacity of general intelligence
Reliability	-
Validity	-
Reference	Kopiez, R., & Lee, J. I. (2008). Towards a general model of skills involved in sight reading music. *Music Education Research, 10*(1), 41-62.
See Also	Kopez, R., Weihs, C., Ligges, U., & Lee, J. I. (2006). Classification of high and low achievers in a music sight-reading task. *Psychology of Music, 34*(1), 5-26.

See Also	Kopiez, R. & Lee, J. I. (2006). Towards a dynamic model of skills involved in sight reading music. *Music Education Research, 8*(1), 97-120.
See Also	Lee, J. I. (2004). *Component skills involved in sight reading music*. Frankfurt am Main: Peter Lang.

Title	Classification of high and low achievers in a music sight-reading task
Author	Reinhard Kopiez, Claus Weihs, Uwe Ligges, Ji In Lee
Year	2006
Purpose	• To classify sight reading performance for the first time. • To extend the expertise approach and test the hypothesis that sight reading expertise is a necessary predictor of sight reading performance
Measurement	• General cognitive skills (e.g. short-term and working memory) • Elementary cognitive and psychomotor skills (e.g. simple reaction time and speed of information processing) • Practice-related skills (e.g. general piano expertise, inner hearing ability and accumulated hours of sight reading expertise) • Lateralization (the connection between the two hemispheres of the brain), and inner hearing (ability to imagine the sound of a score)
Assessments Used	-
Description	
Subjects	52 piano university students (mean age = 24.56). Their majors were eight piano, or they had to be experts in piano chamber music or accompanying
Instruments Used	Piano
Stimuli	Pre-recorded pacing melody, with two warm-up pieces and 5 pieces of increasing complexity. Taken from sight reading literature (UNISA), but rearranged for solo melody and piano accompaniment. Pre-recorded solo part done by violin
Apparatus	*MIDI* piano for performance *Cubase* for recorded solo melody *MidiCompare* (Dixon, 2002) for scoring
Method	See Lee (2004)
Results	Able to demonstrate that classification of sight reading performance is useful for sight reading achievement analysis. Results showed that sight reading achievement for high level pianists is determined by acquired expertise as well as by invariant factors (e.g. speed of information processing and psychomotor speed)
Reliability	-
Validity	-
Reference	Kopiez, R., Weihs, C., Ligges, U., & Lee, J. I. (2006). Classification of high and low achievers in a music sight-reading task. *Psychology of Music, 34*(1), 5-26.
See Also	Kopiez, R. & Lee, J. I. (2006). Towards a dynamic model of skills involved in sight reading music. *Music Education Research, 8*(1), 97-120.
See Also	Kopiez, R., & Lee, J. I. (2008). Towards a general model of skills involved in sight reading music. *Music Education Research, 10*(1), 41-62.
See Also	Lee, J. I. (2004). *Component skills involved in sight reading music*. Frankfurt am Main: Peter Lang.

Title	An exploratory study of individual difference variables in piano sight-reading achievement
Author	Ladonna Eloise Kornicke
Year	1992
Purpose	• To examine selected individual difference variables as predictors of piano sight-reading achievement, including sight-reading experience, aural imagery, locus of control, field dependence/independence, and personality variables. • To determine the best predictors among the respective individual difference variables and sight-reading achievement. Primary Research Questions: 1. What is the best linear combination of the predictor variables and sight-reading achievement? 2. What is the best combination of predictor variables distinguishing between the extreme groups of high and low sight-reading achievers?
Measurement	• Test of Aural Imagery (AI) for subjects' ability to compare/contrast an aural musical example with a printed notational version • Hidden Figures Test (FDI) for subjects' relative Field Dependence/Independence (FD/FI) • Locus of Control Test (LOC) for subjects' attribution of success/failure to internal or external sources • Sight-Reading Experience Questionnaire (SREQ) for subjects' sight-reading experience • Myers-Briggs Type Indicator for personality, measuring: extraversion/introversion (EI), sensing/intuitive (SN), thinking/feeling (TF), judgment/perception (JP) **Sight Reading Task:** A Sight-Reading Performance Scale (SRPS) of 26 items (later revised to 32 items) was developed to adjudicate the taped performances for each of the 5 pieces. Each item was scored on a 5-point scale Judges for SRAT were doctoral level students with a completed Masters degree in piano performance. They previously studied the questions and scored the performances independently. They also received training **Sight-Reading Experience Questionnaire:** Some items were scored from 1-5, and treated as individual variables. Other items were scored by combining the 6 types of sight-reading experiences **AI:** 1 point for each chord, for a grand total of possible 90 points **MBTI:** Scored according to published directions FD/FI: Scored by counting number of correct responses (possible range was 0-32. Low scores = relative field dependence. High scores = relative field independence)
Assessments Used	Myers-Briggs Type Indicator (MBTI), Form F for measuring personality variables (Myers & Briggs, 1976) Researcher-constructed sight-reading test (SRAT): Comprised of likert-scale ratings Hidden Figures Test (HFT) (French, Eckstrom, & Price, 1962) for cognitive style of field dependence/independence
Description	
Subjects	73 subjects; a cross-section of college-level pianists from 4 state universities in Indiana and Illinois, as well as piano faculty members. Population limited to undergraduate, graduate, or professional pianists
Instruments Used	Piano
Stimuli	**Sight-Reading:** 5 excerpts, 2 pages in length, without title or composer. Chosen by researcher from available concert repertoire: Scriabin, Prelude Op. 13, No. 6 b minor; Bach, Capriccio in Bb major; Lajos Papp, Bagatelle No. 5; Beethoven, Twelve Variations in A major; George Perle, Etude No. 1. **Sight-Reading Experience Questionnaire:** 16 questions, divided into 3 sections: pre-college experience, college experience, and knowledge of musical style, with 3 response formats: categorical, open-ended response, and likert scale ratings **Aural Imagery Test:** Recorded chords on acoustic piano. Part 1 was 30 individual

	chords **Locus of Control Test:** 46 items that employed 5-point likert scales. **FD/FI:** Hidden Figures Test: 32 examples and practice section, to identify which of 5 simple geometric figures was embedded in each complex geometric figure
Apparatus	Tape-recorder for performances Video Tape and Television Screen for Aural Imagery test Stop Watch Pencil and paper
Method	The SRAT, AI, SREQ were in a pilot study before the main study For Main study, 2-hr appointments were made for each of the 73 subjects. Tests administered in this order: SRAT, SREQ, AIT, LOC, MBTI, FDI. Before SRAT, subject recorded their ID number on audio-tape. They were instructed regarding sight-reading procedures. For each piece, allowed 30 s to study it, timed with stop watch. **SRAT:** Subjects took the SRAT in two phases: Phase 1 – audio-taped interviews with 12 pianists (Kornicke, 1989). Phase 2 – 5 musical examples. Subjects set their own tempo, and had to play the entire excerpt. **AIT:** Section A: Item number read, silence, visual stimulus projected for 1 s, silence, aural stimulus sounded for 2 s, silence for 5 s to mark on the response sheet. Subjects were to determine if visually displayed chord matched the aurally sounded chord. Section B: Item number read, silence, visual stimulus projected for 2 s, aural stimulus sounded for 3 s. First 2 chords written as half-notes sound for 1 s, and final chord written as whole note for 2 s, silence for 6 s to mark on the response sheet. Subjects saw three chords in this section, with the knowledge that one of the latter two could be altered **MBTI:** Form F. 30-45 min to complete. Administered according to published directions **FD/FI:** Hidden Figures Test: 15 minutes for test
Results	Found that selected individual differences variables served as significant predictors of sight-reading achievement. Found that Aural Imagery, Sight-reading Experience, and Field Independence were best linear combination of predictor variables of sight-reading achievement for entire sample. Found that best combination of predictor variables distinguishing between high and low sight-reading achievers was SR-EXP, FI, AI, and thinking from the MBTI. Also found, though not as primary, 1. FI and external LOC related to sight-reading achievement for males, 2. Suggestion that subjects' ability to sight-read not contingent upon a particular style of music but on general ability to sight-read across styles, 3. Strongest relationships between SR-EXP and item analysis of SRAT were for items related to interpretation, 4. Strongest correlations between FI and SRAT were for items related to pitch accuracy, 5. External LOC was correlated with several items related to interpretation, but not to tempo or pitch items, 6. Several personality items were correlated with SRAT, including thinking with pitch accuracy, intuition with tempo maintenance, introversion with shaping musical lines, perception with accurately performing rhythmic notation, 7. Sight-reading experience included factors related to quantity, frequency, range of sight-reading experience in addition to factors of attitude and initiative toward sight-reading, 8. Number of pieces performed was not related to sight-reading achievement, 9. Moderate correlation between beginning sight-reading at a later age and sight-reading achievement
Reliability	**Sight-Reading Performance Scale (for SRAT):** Inter-judge reliability for composite for pilot study = .93; for main study = .99
Validity	-
Reference	Kornicke, L. E. (1992). *An exploratory study of individual difference variables in piano sight-reading achievement* (Order No. 9301458). Available from ProQuest Dissertations & Theses Global. (304007507). Retrieved from http://search.proquest.com/docview/304007507?accountid=14701

Title	The effects of error-detection practice on keyboard sight-reading achievement of undergraduate music majors
Author	Marilyn J. Kostka
Year	2000
Purpose	• To gain further insight into the relationship between error detection and sight-reading ability • To measure the effects of error-detection practice on keyboard sight-reading skills
Measurement	• Errors in sight-reading tests Scored by 2 independent observers with combined piano-teaching experience of 30+ years, listening to audiotapes while completing a form for note errors, rhythm errors, and hesitations. Scored both pretest and posttest. Errors were counted according to whether they occurred in a specific measure
Assessments Used	Pretest: Sight reading test
Description	
Subjects	69 undergraduate music majors enrolled in 6 intact keyboard classes: 3 classes in their first semester of instruction and 3 classes in their third semester of instruction. One class form each level was assigned to one of three experimental conditions: 1. Error-detection practice plus shadowing (EDS), 2. Shadowing only (S), 3. Self-guided practice (CC), and to one of the two instructors
Instruments Used	Piano
Stimuli	Pretest: 2 pieces (subjects played just one) appropriate for Level I or Level III, 8 measures in length and 4/4 time Posttest: In the 15th week, individual sight-reading test administered to all subjects. It was identical to the pretest except for a more difficult sight-reading piece to account for gains in skill
Apparatus	*Kurzweil Mark-10* electronic keyboard for performances *TEAC* tape player with headphones for subject and instructor, and for recording
Method	Pretest: Subjects looked at piece for 15 s, then played to best of their ability with goal of maintaining chosen tempo For 10 weeks, all classes had 5 brief sight-reading sessions (5-7 minutes) every 2 weeks EDS Group: Before sight-reading, they had 5 minutes of practice in error detection, which was listening to 3 repeated examples of same prerecorded piano piece while visually following score and identifying mistakes form the recording by marking score. The instructor discussed errors with the class and gave correct answers EDS and S Groups: They were instructed to shadow-practice music before playing CC Group: Told to practice sight-reading on their own using any method
Results	No significant differences found between groups. All groups improved in rhythm and accuracy, but combined frequency of hesitations remained nearly same form pretest to posttest. Data suggests that error-detection practice may result in overall improved sight-reading achievement
Reliability	Inter-rater reliability = .92
Validity	-
Reference	Kostka, M. (2000). The effects of error-detection practice on keyboard sight-reading achievement of undergraduate music majors. *Journal of Research in Music Education, 48*(2), 114-122.

L

Title	Component skills involved in sight read music
Author	Ji In Lee
Year	2004
Purpose	• To define and test the component skills involved in sight reading, using 3 categories of tests: General cognitive skills, Elementary cognitive skills, Practice skills. Study is based on hypothesis that expertise is a necessity but not a sufficient predictor for sight reading achievement. Research questions: What is the role of expertise in relations to other factors; and which factors other than sight reading expertise can predict sight reading achievement? What is the role of actors related to cognitive capacity in sight reading? What is the role of working memory and short-term memory in sight reading? What is the role of elementary cognitive skills such as mental speed in sight reading? What is the role of inner hearing in sight reading? Which combination of component skills can be a predictor for sight reading achievement? Which factors influence sight reading ability as the level o complexity increases?
Measurement	• Measure and compare component skills of sight reading. (With "divergence of component skills, one can hope to obtain variances which can be explained independently from each other." P. 24) • Tapping Speed: Measured tapping speed of both hands over 30 s. Median values of inter-tap interval (ITI) over 30 s. Measured median number of taps per s (in Hz), for each hand • Trilling Speed: Measured speed of psychomotoric optimization and its possible correlation to sigh reading achievement • Short-tem (music-specific) memory test: Recorded performances, and a point was given for each correct note • Raven's D Matrice: Measure some type of reasoning ability. In test manual, states that it measures the general ability to perceive and think clearly, and educative ability • Sight Reading Performance: Used Midi-Compare to analyze performance. Number of correct notes for both hands was calculated for window of time for when a note is supposed to be played • Short-Term Music-Specific Test: Printed performance on sheet and circled correct notes and counted them. One point for each correct note. Only measured correct pitches. An external examiner also scored this for an objective opinion • Retrospective Interview Questions: Calculated hours with specific questions
Assessments Used	Lehmann & Ericsson (1996) questionnaire *Raven's D Matrice* (Intelligence) (J.C. Raven, 1938) *Number connection test* (*Zahlen-Verbindungs-Test*) (Oswald & Roth, 1997)
Description	
Subjects	52 pianists: piano major students or piano graduate students for Hanover University of Music and Drama (HMT) (mean age = 24.56 years)
Instruments Used	Piano
Stimuli	**Sight reading:** Stimuli divided into 5 different levels of increasing complexity, with 2 warm-up pieces. Stimuli had pre-recorded pacing melody to put constraints on subjects to perform in real time and avoid tempo changes. Stimuli were unknown and similar lengths, with a recomposed solo line added. Two lists given to 2 evaluators - List A: Stimuli taken from Four Star books (The Royal Conservatory of Music [Toronto, Canada]); books 8-10, List B: Stimuli taken from University of South Africa (UNISA) piano sight reading syllabus book; grades 5 – 8. Each list had 15 pieces; 3 for each level of difficulty. Evaluators chose List B for this study and pieces were rewritten by a composer, with solo voice recorded by a professional violinist. **Questionnaire:** List of questions selected from Lehmann and Ericsson (1996).

	Questions concerned accumulated hours of practice and performance of sight reading, accompaniment and solo practice. **Audiation Test:** 21 themes and variations initially chosen by author from pre-existing music literature (by Verdi, Beethoven, Mozart, Sweelinck) Composer chose 15 of these and composed a 'lure melody' for each. Lure melody is similar to the theme but distinctively difference and has significant deviation from underlying melodic structure of the melody. For each example, there would be 3 versions – theme, variation and lure melody. Finally, they chose 2 warm-ups and 5 samples from original 15. **Reaction Time Test (RT):** Used researcher-developed software. **Number connect test (NCT):** Paper has shortened version of many numbers arranged on paper for subjects to connect in order. **Short-tem music memory AND Short-term numerical memory:** These were taken from Oberaurer et al. (2000) test battery on working memory capacity. **Short-term (non-music-specific) numerical memory test:** 4 digits per ask, increasing by one digit every time until it was 9 digits per task. **Short-term (music-specific) memory test:** 3 single-line melodies.
Apparatus	*FINALE* for writing of sight reading pieces *MIDI* keyboard to record performances of subjects *CUBASE (VST) 5.1* (Steinberg, 2000) for sequencing of performances, and for printing the STMM test results Pencil for questionnaire *PowerPoint* presentation method for audiation test Speakers for playing of recording for audiation test Morse code key PC in DOS-Mode for tapping test Computer screen for presentation of stimuli Stopwatch for examiner *Midi-Compare* (Dixon, 2002) for analysis of sight reading performance
Method	**Order of Test Battery:** Subjects entered lab, introduced to the 3 researchers and lab set-up. Total time 2.5-3 hours/subject. *Summary* *Section 1* Recording of Personal details 1. Number connection test 2. Reaction time test: a. Visual (7 warm-ups & 15 tasks), b. Auditory (7 warm-ups & 15 tasks) 3. Tapping for 30 s: RH, LH 4. Short-term non-music-specific memory test (2 warm-ups & 18 tasks) 5. Working memory test (6 warm-ups & 18 tasks) (This took 1 hour, followed by 15 min break with EEG pads placed on skin.) *Section 2* 6. Sight reading task (2 warm-ups & 5 pieces) 7. Trilling for 15 s: 1st & 3rd fingers (twice), 3rd & 4th fingers (twice) 8. Eye-movement tests (Relevant for this thesis: Raven's Matrice D) (This took 1 hour, following by 15 min break with EEG pads removed) *Section 3* 9. Inner hearing test (2 warm-ups & 5 tasks) 10. Short-term music-specific memory test (1 warm-up & 3 tasks) 11. Retrospective interview (This took 30 minutes) **Sight Reading:** Warm up pieces to begin. Subjects looked at each piece for 60 s. Tempo indicated with 2 full bars of clicks before the piece. Subjects read all stimuli from level 1-level 5 and had to accompany the solo violin with each piece. *General Cognitive Skills* **Short-tem (music-specific) memory test:** Melody line shown to subjects for 1

	minute and then they played as much as they could remember. **Short-term (non-music-specific) numerical memory test:** Subjects were shown series of numbers, one at a time, and were required to remember it. They had to recall the numbers aloud. These responses shown on a screen and subjects could correct their answers if they wanted to. **Number Connection Test (NCT):** Subjects connect numbers for 1-90, which are arranged in such a way that subjects cannot draw straight line, but look for next number. **Raven's D Matrice:** Subjects select best item out of 6-8 pattern pieces, to fit 36 actual items. *Elementary Cognitive Skills* **Speed tapping:** Subjects used fore and middle fingers with their forearms touching the piece of carpet attached to the Morse code key. Subjects chose with which hand to begin. **Simple Reaction time (Auditory and Visual):** Visual - Subjects pressed Morse tapping key when visual stimuli (big circle) appeared on computer screen. Auditory – Same as visual except for auditory stimulus (beep) as cue. Subjects used their preferred hand's fore and middle fingers. **Speed Trill Playing:** Subjects trilled for 15 s using thumb and middle finger of RH on C and D, and then repeated. Then trilled using middle finger and ring finger of RH on E and F, and repeated. *Practice Skills* **Audiation (Inner Hearing) Test:** 2 warm-ups, then 5 examples. Variation of each theme shown for 45 seconds and theme or lure melody heard through speakers. This could be repeated. Followed by a same-different forced choice paradigm. Subjects required to state if theme was embedded in variation or not, while instructor completed answer sheet.
Results	General: Found that the best combination of predictors for sight reading achievement for all 5 levels of difficulty in total are: 1. Speed trill 3-4-2, 2. Amount of sight reading experience gained up to the age of 15, 3. Time taken to complete the Number Connection Test and, 4. Time taken to complete the Speed Trill 3-4-1. (MC's Note: Pg. 80 also reads: "Speed trill, S. R. expertise up to 10 years of age, Number connection tests and Inner hearing test have the highest correlation with total sight reading performance.") Sight reading achievement can be predicted by combination of expertise and general cognitive skills. Expertise is not the only predictor for S. R. achievement but a combination of general cognitive skills, expertise and speed factors required.
Reliability	Advanced Progressive Matrices 9APM) (J.C. Raven, 1947) = .91 Short-Term Music-Specific Test: Inter-rater reliability was 95%, or r(21) = .975
Validity	-
Reference	Lee, J. I. (2004). *Component skills involved in sight reading music*. Frankfurt am Main: Peter Lang.
See Also	Kopiez, R., Weihs, C., Ligges, U., & Lee, J. I. (2006). Classification of high and low achievers in a music sight-reading task. *Psychology of Music, 34*(1), 5-26.
See Also	Kopiez, R. & Lee, J. I. (2006). Towards a dynamic model of skills involved in sight reading music. *Music Education Research, 8*(1), 97-120.
See Also	Kopiez, R., & Lee, J. I. (2008). Towards a general model of skills involved in sight reading music. *Music Education Research, 10*(1), 41-62.

Title	Sight-reading ability of expert pianists in the context of piano accompanying
Author	Andreas C. Lehmann and K. Anders Ericsson
Year	1993
Purpose	• To measure individual differences in sight reading ability
Measurement	Three aspects being measured: • Scoring of pretest sight reading • Comparison of individual performances with their professional specialization • The effect of multiple performances of the same piece for sight reading
Assessments Used	Pretest: Biographical background interview
Description	
Subjects	16 expert pianists; students at Florida State University School of Music. 8 were focused on accompanying and 8 were focused on performance
Instruments Used	Piano
Stimuli	Stimuli and warm-up piece excerpts were from moderately difficult pieces for solo instrument and accompaniment. Warm-Up: F. Schubert, Sonatina in D Major, op. 137, mm. 1-72 MM 80. Test Piece 1: M. Blavet, Sonata in a minor, op. 26, mm. 1-19 MM 60. Test Piece 2: W. Molique, Polonaise in D Major, op. 43, mm. 1-34 MM 80. Each excerpt was two pages in length. 2 measures of metronome click in eighth notes preceded each example
Apparatus	MIDI files for pre-recorded flute tracks, played on an *IBM* computer Keyboard and pedal recorded by *MIDI*
Method	Subjects accompanied two pre-recorded flute parts from a music score with the goal of accurate pitch, rhythm, and expression. Each accompaniment was performed 4 times; the first and last time with pacing, the second and third times without pacing or flute track
Results	With expert musicians, S/R ability does not increase with instrumental skill, but with specialization in accompanying. The findings suggest that expert pianists differ in specialized abilities
Reliability	-
Validity	-
Reference	Lehmann, A. C., & Ericsson, K. A. (1993). Sight-reading ability of expert pianists in the context of piano accompanying. *Psychomusicology, 12*, 182-195.

Title	Performance without preparation: Structure and acquisition of expert sight-reading accompanying performance
Author	Andreas C. Lehmann and K. Anders Ericsson
Year	1996
Purpose	• To identify the source of individual differences in sight-reading ability
Measurement	• Analyze the component task performance and its relation to expert sight-reading and accompanying • Examine biographical data regarding their skill acquisition • Data of acquisition of sight-reading skill Recorded performances were transcribed with software. Each correct note, within a sixteenth note, was scored. For recall and improvisation, only the exact note matches between the original and performance versions for critical passages were scored. The leap task was scored by adding the correctly performed target notes separately for the two performances
Assessments Used	-
Description	
Subjects	16 expert pianists; students at Florida State University School of Music. 8 were focused on accompanying and 8 were focused on performance
Instruments Used	Piano
Stimuli	The first stimulus was a page from the previous study's excerpts from moderately difficult pieces for solo instrument and accompaniment. The page had sections of the

	piece missing. The second stimulus was a page of unfamiliar music with sections missing. The third stimulus was a simple chord progression with large leaps in one or both hands, and was 3 lines long
Apparatus	Computer to record key-strokes Goggles with the lower half of he frame painted black and two 0.5 inch, black, tubular extensions attached to the lenses
Method	Subjects were unexpectedly given 2 tasks regarding selected component skills. The first task – a recall task, where they were given one page of a piece they had played and informed they would play again with pacing. There were sections of the accompaniment missing, and subjects were asked to play whatever they could remember (they were given an opportunity to play through the page for memorization). The second task – an improvisation task, where they also played for a paced flute part, but with an unfamiliar piece, and asked to improvise in the empty sections. The subjects played this twice. The final component task – the leap task, where they would have large leaps in one hand, or both hands. They initially learned the underlying chord progression, and then played the piece twice while paced (MM eighth note = 70). The second time they played this they wore goggles to prevent peripheral vision
Results	The analyses of combined and separate scores confirm that specialized training (accompanying experience) and relevant knowledge (accompanying repertoire) were significant predictors for S/R and accompanying ability. So, individual differences in the ability to S/R and accompany are significantly correlated with the amount of accompanying and available repertoire. The specialized experience and accompanying activities were better predictors of S/R than accumulated piano practice
Reliability	-
Validity	-
Reference	Lehmann, A. C., & Ericsson, K. A. (1996). Performance without preparation: Structure and acquisition of expert sight-reading accompanying performance. *Psychomusicology, 15*(1-2), 1-29.

Title	Music readers and notation: Investigation of an interactive model of rhythm reading
Author	Katherine M. Levy
Year	2001
Purpose	• To examine the rhythmic structuring relationship to music reading through a study of rhythm reading miscues • To explore potential connections between music perception, music's mental representation, and music reading, to improve learning of an instruction in musical rhythm reading Total of 5 research questions: The first 4 investigated the connection between rhythm cognition, perception, and knowledge, and music reading proficiency. The 5th addressed usefulness of the Rhythm Reading Interactive Model and music reading miscue analysis
Measurement	The procedure includes methods for a. defining and coding miscues, b. applying to each miscue model-based criteria discrimination cue system sensitivity (c.f., high, medium or low syntactic acceptability), c. recognizing patterns of cue system sensitivity, d. creating individualized reader profiles, and, e. interpreting reader profiles in terms of reading proficiency and the learning environment. Measurement analyzed by three judges: The author and two others. In sightreading portion, miscues were scored by: substitutions, omissions, insertions, repetitions, including attempted or successful corrections, and pauses.
Assessments Used	Research Question 4: Reading Interview, Notation Practices and Training Questionnaire, Musical Achievement Test (Colwell 1969-1970, Test 2, Parts 3a and b)
Description	

Subjects	36 grade 7 students (second year of instruction) from a middle school band, though investigated miscues of just 11 participants
Instruments Used	Wind and brass band instruments
Stimuli	Research Questions 1-3: 2-4 etudes in standard notation from band instrumental music instruction course for beginners (Lautzenheiser, Higgins, Menghini, Rhodes, & Bierschenk, 1999) and intermediate students (Lautzenheiser, Higgins, Menghini, Rhodes, & Bierschenk, 2000)
Apparatus	*Finale* for transcription of stimuli Videotape and recorder for recording sessions
Method	1. Readers completed self-reports of their training histories (questionnaires) before sightreading sessions 2. Readers participated in a reading interview with 10 questions: When you are reading music and come to a rhythm you don't know, what do you do? Who is a good rhythm reader you know? What makes (person's name) a good rhythm reader? Do you think (person's name) ever comes to a rhythm she doesn't know? If the answer to question 4 was yes, then ask what the reader thinks (person's name) does when he comes to a rhythm he doesn't know. If the answer to question 4 was no, ask the reader to suppose that comes to something she doesn't know. If you knew someone was having trouble reading rhythms, how would you help that person? What would your teacher do to help that person? How did you learn to read music? What would you like to do better when you read rhythms? Do you think you are a good rhythm reader? Why? 3. The participants completed a Notation Practices and Training Questionnaire for information about readers' training and non-musical knowledge and their preferences for musical activities that included or excluded music reading 4. Two Music Achievement Tests (Colwell, 1969-1970) subtests were selected to provide non-performance measures of readers' auditory-visual discrimination. These were completed by sight or by listening. 5. Field observations of students by researcher in their regular band class for 3 weeks. 6. Each participant sightread a series of etudes with their instrument
Results	Found that from all data the readers' performance-rated ability, rhythm reading proficiency, and reading approaches were associated with patterns of cue and strategy use, as well as with musical training and skills. All readers produced miscues consistent with RRIM's proposed systems. Found that low proficiency readers had few miscues that reflected functioning cue systems, were less sensitive to aural cues, verbalized incomplete or confused conceptual musical knowledge, had difficulty discussing rhythm versus other reading components. Found that proficient readers had higher patterns of Metric and Aural cue use and easily discussed rhythm separately from other notation components. Results also revealed two reading approaches: inventive (perceptual, intuitive) and conventional (training-related). Inventive readers used more figurative pattern cues. Conventional readers' cue use was evenly distributed among patterns
Reliability	-
Validity	-
Reference	Levy, K. (2001). *Music readers and notation: Investigation of an interactive model of rhythm reading* (UMI No. 3034120).

Title	Evaluation of a sight-reading test administered to freshman piano classes
Author	Jerry E. Lowder
Year	1973
Purpose	• To provide an analysis of errors performed on a sight-reading test administered to freshman piano classes at the end of the first semester of instruction • To determine whether there would be any differences between the performances of experimental and control groups • To determine whether the teaching of vertical intervallic relationships according to figured bass principles would improve the pianists' ability to sight-read music based on tertian harmony
Measurement	• Sight-reading proficiency, with specific note of pitch accuracy, steady tempo, rhythmic accuracy, and selection of accompaniment chords. However, only pitch accuracy is analyzed in this study Three experienced class piano teachers from University music faculty evaluated recordings
Assessments Used	Author-composed sight-reading test, administered at the end of the first semester
Description	
Subjects	23 subjects: 22 freshman music majors and 1 freshman non-music major
Instruments Used	Piano
Stimuli	Author-composed or author-selected, and typical of the materials presented during the subjects' first semester of piano class. 8 of the 15 musical examples are used in this study for assessment
Apparatus	Tape recorder for performances
Method	Control and experimental groups received the same instruction in sight-reading and other keyboard skills. The experimental group also used a program that stressed reading and performance skills, emphasizing intervallic relationships and figured bass
Results	They found no significant differences in pitch accuracy between the two groups With analysis of pitch errors: Found that a pitch error usually was accompanied by a rhythmic error Found that a high proportion of subjects' errors consisted of omission of accidentals Found that more errors occurred in bass clef than in treble clef
Reliability	-
Validity	-
Reference	Lowder, J. (1973). Evaluation of a sight-reading test administered to freshman piano classes. *Journal of Research in Music Education, 27*(1), 68-73.

M

Title	Music reading ability of beginning wind instrumentalists after melodic instruction
Author	Carol B. MacKnight
Year	1975
Purpose	• To research and develop teaching techniques and materials to treat music reading as a process and emphasize the structure of the melodic line • To determine if there is any difference between scores of experimental and control groups: On Watkins-Farnum Performance Scale (WFPS)? For brass or woodwind instruments on WFPS? For high or low musical aptitude on WFPS? On Music Achievement Test (MAT)? For brass or woodwind instruments on MAT? For high or low musical aptitude on MAT? For concept scores of each group on pretest and posttest administration of Student Attitude Questionnaire (SAQ)?
Measurement	• WFPS: Errors of pitch, time, change of time, expression, slurs, rests, holds, pauses, repeats • MAT: Major-minor mode discrimination of chords and phrases; feeling for tonal center within cadences and phrases; auditory-visual discrimination of pitch and rhythm • SAQ: Author-composed for observing and measuring the psychological meaning of concepts
Assessments Used	Pretest: Musical Aptitude Profile; Student Attitude Questionnaire; Recent scores from Lorge-Thorndike Intelligence Test Posttest: WFPS (Form A); MAT (Test 2); Student Attitude Questionnaire
Description	
Subjects	Approximately 90 grade 4 students, all woodwind players, without private lessons during the experiment. Grouped into classes of 6 students, in either experimental group or control group
Instruments Used	Voice (for training) and wind instruments
Stimuli	Experimental Group: Author-prepared instructional book with melodies containing the tonal patterns, in same order of presentation Control Group: *Breeze Easy* method book, following exercises in order
Apparatus	-
Method	Weekly 30-min of instruction for 32 weeks. The method of introducing pitches: Experimental – Introduce with series of tonal patterns using solfege. Each pattern was taught in 3 stages: 1. Aural presentation, 2. Auditory-visual presentation, 3. Auditory-visual presentation of pattern within a phrase. Students responded vocally and instrumentally. Rhythm was also presented in phrases, using syllables for the values. Control – Used method book, but also used rhythm syllables and sang assigned melodies with letter names
Results	Suggest that melodic training based on auditory-visual recognition of tonal patterns is effective for teaching notation to beginning readers. This method is superior with development of sigh treading and auditory-visual discrimination skills. Instruction should emphasize 1. Identification of patterns, 2. Active listening, 3. Singing tonal syllables, 4. Changing rhythm syllables, 5. Thought and conceptualization, 6. Materials that introduce tones and rhythms in their frequent patterns
Reliability	WFPS = .93 for all tests for musically select fourth-grade student groups MAT = .91 SAQ = .81
Validity	WFPS: Construct validity coefficients ranging from .64 to .94 with median of .79 for performance groups

	MAT = .92
Reference	MacKnight, C. B. (1975). Music reading ability of beginning wind instrumentalists after melodic instruction. *Journal of Research in Music Education, 23*(1), 23-34.

Title	Factors and abilities influencing sightreading skill in music
Author	Gary E. McPherson
Year	1994
Purpose	• To replicate and extend existing literature • To determine important factors and abilities that influence sightreading skill in music
Measurement	• WFPS: Sight reading ability, with categories of pitch, rhythm, slurring/articulation, tempo, expression, pause/fermata, repeats (used for unrehearsed music) • AMEB: Tone quality, intonation, musical styles and interpretation, phrasing, ability to keep in tie with an accompaniment (used for rehearsed music) Author-scored WFPS during each research session Errors notes with abbreviated system of symbols above each measure
Assessments Used	Watkins-Farnum Performance Scale, Form A (Watkins & Farnum, 1954) Australian Music Examinations Board (AMEB, 1990)
Description	
Subjects	101 high school instrumentalists, all preparing for an Australian Music Examinations Board (AMEB) performance examination, from a wide area, undertaking grades 3 or 4 (Group 1, average age = 13 years), or grades 5 or 6 (Group 2, average age = 16 years, 1 month)
Instruments Used	Clarinet or trumpet
Stimuli	WFPS: 14 increasingly-difficult sight reading exercises
Apparatus	*Sony Walkman Professional* stereo cassette-recorder (*WM-D6C*) to record performances Metronome to establish tempo for WFPS
Method	Each subject completed Form A of WFPS, as best as they could, with 20 s between each exercise
Results	Group 1 subjects with high result at AMEB exam were not the best sight readers Group 2 strong sight readers were more likely to have received a high award in the AMEB. The weakest readers did not take good note of important features (e.g. key signature). The strongest readers were cognizant of important details (e.g. time signature). Many of the poor sight readers appeared to be unable to process the music they were performing. In contrast, the strong readers appeared to anticipate the flow of the music
Reliability	Based on pilot study, inter-judge reliability = .99 for 2 separate scorings of WFPS by researcher, and .98 for the 2 independent judges
Validity	-
Reference	McPherson, G. E. (1994). Factors and abilities influencing sightreading skill in music. *Journal of Research in Music Education, 42*(3), 217-231.

Title	The assessment of music performance: Development and validation of five new measures
Author	Gary E. McPherson
Year	1995
Purpose	• To examine the ability of high school instrumentalists to perform a variety of performance skills • To develop and validate 3 new measures of high school instrumentalists' ability to play from memory, play by ear, and improvise • To compare these above abilities with results obtained on existing measures designed to asses subjects' ability to sight-read and perform rehearsed music

	• To determine the degree of correlation between 16 variables related to subjects' musical background and five performance skills: ability to sight-read, perform rehearsed music, play from memory, play by ear, and improvise
Measurement	• Questionnaire to obtain information on 16 variables believed to influence the development of the above skills Tests: • Sight-Reading and Performing Rehearsed Music: Administered by "specialist examiners" using a 12-point scale. Scale ranged from Grade 3 "Satisfactory" to Grade 6 "Honours" Playing Music from Memory, by Ear, and by Improvising: After interview with professional musicians, the measure went under a period of field-testing and analysis with a pilot test of 45 high school students, then refined for use in this study. Three judges worked over a period of 5 months to assess a total of 5,722 responses on the 3 new measures (see McPherson, 1993a) • Test of Ability to Play from Memory (TAPFM): A practice item and 4 test items. 3 judges were trained, using the practice item, with scoring methods. For Part 1, judges scored from 0-5, that best represented their evaluation of the performance. Their assessment was based on musical accuracy in pitch, rhythm, and phrasing. For Part 2, they scored from 0-7 • Test of Ability to Play by Ear (TAPE): For the two familiar pieces, evaluators rated each measure as correct, partially correct, or incorrect. They marked respectively with "0, 1, or 2". As subjects performed each example and transposition twice, evaluators scored each attempt separately. Score was composite mark for first and second performances in both the original and transposed key • Test of Ability to Improvise (TAI): Newly devised and incorporated aspects of measurement of creativity in music and skills thought essential by researchers and expert improvisers (See McPherson 1993a, 1993b). Three judges scored these independently, with four 5-point rating scales: Instrumental Fluency (Hesitant and labored to Spontaneous and confident), Musical Syntax (Illogical to Logical), Creativity (No uniqueness to Marked uniqueness), Musical Quality (Unappealing to Appealing). Judges listened to the recordings and scored the first two dimensions with the first listen, and the last two dimensions with the second listen • Questionnaire: 1. Early exposure to music - 3-point scale, 2. Starting method on instrument - Dichotomous variable: 1 or 0, 3. Total period of formal training on the musical instruments - Reported in months from 14-134, 4. Period playing clarinet or trumpet - Reported in months from 14-108, 5. Period of private lessons - Reported in months from 0-102, 6. Learning AMEB theory/musicianship - 1 or 0, 7. Learning other instrument(s) - 3-point scale, 8. Average daily practice - Average over previous month, between 2-90 minutes, 9. Electing School Music - 0 or 1, 10. Participation in ensembles - Report of number of ensembles on regular basis; 0-4, 11. Report of frequency of learning pieces from memory - Reported as "often", "sometimes", "rarely", "never"; coded from 1-4, 12. Report of frequency of playing by ear - Same scoring as variable 11, 13. Report of frequency of improvising - Same scoring as variable 11, 14. Report of frequency of composing - Same scoring as variable 11, and aggregated to form a composite score, 15. Report of interest and participation in singing activities - Same scoring as variable 11, and aggregated to form a composite score, 16. Report of frequency of mental rehearsal - Same scoring as variable 11
Assessments Used	Watkins-Farnum Performance Scale (WFPS), Form A (Watkins & Farnum, 1954)
Description	
Subjects	101 high school clarinet and trumpet players, preparing for an AMEB performance examination (AMEB Grades 3-6). They were grouped according to 2 levels of age

	and ability: Group 1 – high school years 7-9, AMEB grades 3-4, ages 12-15; Group 2 – high school years 10-12, AMEB grades 5-6, ages 15-18
Instruments Used	Clarinet and Trumpet
Stimuli	**Sight-Reading and Performing Rehearsed Music:** Form A of the WFPS **TAPFM:** Practice item and 4 test items. Practice item was notated 2-measure melody in C Major. Part 1 was 2 short melodies (F and G Major). Part 2 was 4 measures in a minor with musical expression. Second part of Part 2 was a 4-measure melody in F Major **TAPE:** Two parts. Practice item was the Australian National Anthem. Part 1: *Happy Birthday* (F and G Major) and *For He's a Jolly Good Fellow* (C and F Major). Part 2: Practice example and 4 previously unknown short melodies **TAI:** First two items were *Closing Phrase 1* and *Closing Phrase 2*, which was a 4-measure "question". Item three was *Rhythm* with a given rhythm pattern. Items four and five were *Motif 1* and *Motif 2*, which consisted of the opening phrase to be used. Item six was *Accompaniment*, which was a recorded piano accompaniment. Item seven was *Freely Conceived*, and did not have a stimulus **Questionnaire:** 1. Early exposure to music, 2. Starting method on instrument, 3. Total period of formal training on the musical instruments, 4. Period playing clarinet or trumpet, 5. Period of private lessons, 6. Learning AMEB theory/musicianship, 7. Learning other instrument(s), 8. Average daily practice, 9. Electing School Music, 10. Participation in ensembles, 11. Report of frequency of learning pieces from memory, 12. Report of frequency of playing by ear, 13. Report of frequency of improvising, 14. Report of frequency of composing, 15. Report of interest and participation in singing activities, 16. Report of frequency of mental rehearsal
Apparatus	Tape Recorder for performances and analysis
Method	Individual appointments for each subject in a familiar setting, such as their home or teacher's studio. The sessions involved the measures of the ability to sight-read, play from memory, play by ear and improvise, and complete a questionnaire. **TAPFM:** Subjects studied practice item for 20 s, and then reproduced on their instrument. For two short melodies of Part 1, subjects were allowed a 20 s period for silent study before attempting to reproduce the melody. For the first item of Part 2, subjects studied silently for 1 minute before reproducing it. For final item, each subject rehearsed 4 times on instrument and studied silently between these performances for 15 s. Then they reproduced the melody on their instrument. For all of these, subjects were asked to reproduce the melody twice, even if first time was correct. **TAPE:** Practice activity was performed to familiarize themselves. Part 1: Familiar pieces. Part 2: Practice example and 4 melodies, which subjects heard, played 4 times on their own instrument, with 1-measure gap between. After listening to each performance, they performed each melody as it was heard in the original key. After their second performance, they transposed the melody and performed it twice **TAI:** For first two items, subjects improvised an "answer" to the given "question". For item three, subjects improvised a melody in a set key using the durations of the given rhythmic pattern. For items four and five, subjects were given an opening phrase as the basis for their improvisation. Item six included a pre-recorded piano accompaniment for which the subjects improvised a melody, in the same style. Item seven was an extended improvisation in any styles or mood that the subject chose
Results	There is evidence of the validity of the 3 new measures. The 5 types of performance correlation coefficients give some indication of the interrelatedness with the older group. Found that the skills of playing by ear and improvising were most closely associated with variables which require an ability to "think in sound", including playing from memory, playing by ear, improvising, singing, composing, and mental rehearsal. The results reinforce the 'belief' that early exposure to music is positively related to aural skills needed to perform
Reliability	**TAPFM:** Inter-judge reliability = .87-.97 for the 3 independent judges and = .90-.97

	for composite scores of the 2 performances for each item **TAPE:** Inter-judge correlations = .81-.99 for each of the 4 separate performances on each item, for the 3 independent judges. The composite marks = .93-.99 for each item of the TAPE **TAI:** Inter-judge correlations = .71-.94 for the 3 judges for the 4 separate item assessments, and .89-.97 for composite marks for each individual task
Validity	-
Reference	McPherson, G. E. (1995). The assessment of musical performance: Development and validation of five new measures. *Psychology of Music, 23*(2), 142-161.

Title	Path analysis of a theoretical model to describe the relationship among five types of musical performance
Author	Gary E. McPherson, Michael Bailey, Kenneth E. Sinclair
Year	1997
Purpose	• To examine the interrelationships among five types of musical performance and 16 variables believed related to their development • To test the links hypothesized between the five performance skills and variables believed related to their development
Measurement	Three judges scored the TAPFM, TAPE, and TAI after each session Information from questionnaire was analyzed
Assessments Used	Watkins-Farnum Performance Scale (WFPS) (Watkins & Farnum, 1954) Test of Ability to Play by Ear (TAPE) Test of Ability to Play from Memory (TAPFM) Test of Ability to Improvise (TAI) Author-composed questionnaire
Description	
Subjects	101 high school clarinet and trumpet instrumentalists in grades 7-12, preparing for Australian Music Examinations Board (AMEB) performance examination. Group 1 (n=53) ages 12-15, Group 2 (n=48) ages 10-12.
Instruments Used	Clarinet and trumpet
Stimuli	Author-composed questionnaire: Identification of 16 variables thought to influence the development of the five performance skills, as a result of discussions with 4 researchers who had conducted students in the area (Colwell, 1961; Elliot, 1982; Elliot & May, 1980; Fowler, 1988; Miller, 1988; Prist, 1989; Webster, 1977). Questionnaire included range of variables associated with students' early exposure and experience in music, length and type of training the student had received, enrichment variables such as playing from memory, playing by ear, improvisation in daily practice, singing and audiating. Questionnaire also asked about participation in ensembles, music study at school, whether they had private theory lessons, and learning or learned other instruments
Apparatus	Tape recorder for performances
Method	Subjects participated in each session, completing the questionnaire, and then the various assessments after each session
Results	Found that the ability to perform a repertoire of rehearsed music most heavily influenced by length of study and the ability to sight-read. Found that skill of improvising most strongly influenced by the capacity of musicians to play by ear. Also found a strong path coefficient between the ability to play by ear and the ability to sight-read. Found a total effect between an ability to sight-read and play from memory, and an ability to play by ear and play from memory. Enriching activities and early exposure exert most influence on the skill of playing by ear. Found that less-experienced Group 1 subjects had not yet developed their ability to audiate relatively straightforward music from notation
Reliability	Inter-judge reliability = .87-.97 (TAPFM), .81-.99 (TAPE), .71-.94 (TAI)
Validity	-
Reference	McPherson, G. E., Bailey, M., & Sinclair, K. (1997). Path analysis of a theoretical

	model to describe the relationship among five types of musical performance. *Journal of Research in Music Education, 45*(1), 103-129.

Title	Deliberate practice is necessary but not sufficient to explain individual differences in piano sight-reading skill: The role of working memory capacity
Author	Elizabeth J. Meinz and David Z. Hambrick
Year	2010
Purpose	• To evaluate subjects with a wide range of piano-playing skill (novice to expert) with basic abilities, such as working memory capacity and whether or not it adds to the prediction of expert performance, above and beyond deliberate practice. In other words, does WMC contribute to individual differences in sight-reading skill at high levels of deliberate practice
Measurement	• Main and interactive effects of deliberate practice and working memory capacity on sight-reading performance • Sight Reading • Perpetual speed, with 4 tasks requiring speeded judgments of letter pairs, number pairs, pictures, and patterns (See Hambrick & Oswald, 2005) • Interview: Estimates of deliberate practice and sight-reading practice were multiplied by 52 and summed to created measures of cumulative deliberate practice. Sight Reading: Scored on scale of 1-7 by 2 expert raters, each of who taught university-level piano and held grade degree in music. Rated each performance independently and blindly on 3 dimensions: technical proficiency, musicality, overall performance. (Score of 0 occasionally when piece was too difficult for a particular subject to attempt). WMC: 1 point awarded for each correct response sheet item and each correct judgment in yes/no decision component. Score was total number of points
Assessments Used	-
Description	
Subjects	57 subjects recruited through advertisements, professional contract, and a university music department (M = 30.9 years). 47.4% of subjects held one or more music degrees, with 80.7% considering piano their primary instrument
Instruments Used	Piano
Stimuli	Background Questionnaire: Questions about primary and secondary musical instruments, professional experiences, and lessons. Sight-reading Test: 6 pieces from *Four Star Sight Reading and Ear Tests* series (Berlin & Markow, 2002). Excerpts from Book 6, Book 8, and Book 10 (low, medium, high difficulty, respectively). One piece in each level was chosen for level of technical difficulty and one chosen for the degree to which it allowed artistic expression. The recommended tempo was used for each piece WMC: 1. Equations with a word, 2. Sentences, 3. Rotated letters with arrow, 4. 4x4 arrays of the letter 'L', possibly also with target. Also, a different matrix had one blue cell appeared after letter and target
Apparatus	*Yamaha Portable Grand DGX-500* keyboard Audio-recorder for performances
Method	Subjects tested individually in 3-hr sessions with materials administered in a fixed order: Experience questionnaire and interview, and Sight-reading. After questionnaire, the interview included asking subjects about their piano-playing history. Subjects indicated on a time line when lessons began and ended, when teachers changes, and other events, to provide reference points for estimates (following procedures used by Ericsson, Krampe, & Tesch-Romer (1993)). Subjects estimated for each year the number of hours per week spent on deliberate practice alone, with goal of improving performance. Also estimated number of hours spent on sight-reading practice. For sight-reading, subjects familiarized themselves with electronic piano. They were

	given 60 s to review each piece, then asked to play without stopping or repeating incorrectly performed passages. After first attempt, given another 60 s to review the piece, and play it again. For WMC, there were 4 tasks. Each task had decision component and memory component: Operation Span Task – included an equation with an answer, and a word. Subject said "yes" if answer was correct and "no" if answer was incorrect, and remembered the word. Reading Span Task – Equations replaced with sentences. Subjects responded "yes" or "no" to indicate whether sentences made sense. Rotation Span Task – Rotated letter and arrow pointing away from letter. Subject instructed to say "yes" if rotated letter was oriented normally, o "no" if it was displayed as mirror image, and to remember direction of arrow. Matrix Span Task – Subject instructed to say "yes" if target was present or "no" if it was not. Then a different matrix with one blue cell appeared and subject was to remember its location. For each task, after responding to 3-6 pairs of components, subject was instructed to record the items to be remembered on a response sheet in the order in which they had been presented.
Results	Found that deliberate practice accounted for nearly half of the total variance in piano sight-reading performance. There was an incremental positive effect of WMC, and no evidence that deliberate practice reduced this effect. Evidence indicates that WMC is highly general, stable, and heritable; therefore questioning whether expert performance is solely a reflection of deliberate practice. WMC may limit the ultimate level of performance that can be attained. Authors suggest that deliberate practice – necessary for acquiring expertise – will not always be sufficient to overcome limitations due to basic abilities. Authors speculate that WMC plays direct role in sight-reading performance by determining extent to which pianists can prepare for future keystrokes by looking ahead in music scores
Reliability	Average of overall ratings across raters, attempts, and difficulty levels: WMC = .81, Sight-reading performance = .99
Validity	-
Reference	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.

Title	An assessment of the vertical method and the ensemble approach for teaching sight reading to secondary class piano students
Author	Lynne Howe Micheletti
Year	1980
Purpose	• To investigate the extent to which improvement of proficiency in sight reading was affected by either the vertical method or the ensemble approach of teaching sight reading to class piano students. • To evaluate two approaches for improving sight reading instruction in piano class. • To investigate the relationship between a number of concomitant variables (i.e. number of years of previous piano instruction, number of previous classes in music theory, class level, area of major, major instrument) and improvement in sight reading
Measurement	• Extent of improvement of proficiency in sight reading • Extent of the efficiency (of the methods) • Questionnaire for: number of years of previous piano instruction, number of previous college theory classes, are of major, major instrument, class level • Pre-performance test to assess knowledge of musical terms and symbols • Piano Sight Reading Test adjudicated by 3 graduate students • MAP for tonal imagery, rhythm imagery, musical sensitivity Pre- and post- sight reading tests scored by hand, tabulating correct pitch, duration, dynamic and tempo markings observed (10 weeks between these tests)
Assessments Used	Musical Aptitude Profile (Gordon, 1965) Piano Sight Reading Test, forms A and B (Shake, 1957) for pre- (form A) and post- (form B) training Pre-Performance Test (Shake, 1957) Author-developed questionnaire
Description	
Subjects	40 subjects: 2nd year music majors at the University of Miami. There were 2 experimental groups and 1 control group, from second year, secondary class piano. One experimental group was instructed with the vertical method; the other experimental group used the ensemble approach. The control group used community songs played in unison fashion
Instruments Used	Piano
Stimuli	Control Group: Golden Book of Favorite Songs (1951), a 4-part hymn style community song book Vertical Method Experimental Group: Once-A-Day Sight Reading Program for Piano, Book 1 (Lawrence, 1963) Ensemble Approach Experimental Group: Ensemble Music for Group Piano (Lyke, 1976), with music for 1 piano; 4 hands, 2 pianos; 4 and 8 hands, multiple pianos; 3-6 pianos, and baroque to contemporary classics, with some folk and popular songs
Apparatus	*Biomedical Computer Programs* (BMD, Dixon and Brown, Eds., 1975) for statistical analyses University of Miami *Univac 1100* computer for calculations
Method	Classes met once a week for 50 minutes, for 10 weeks. All classes were taught by the same instructor. Data collected through the Fall semester of classes. The experimental treatments consisted of 15 minutes of guided sight reading once a week. The control group received equal guided sight reading time using community songbook. All students completed demographic questionnaire on first day of classes, and completed the pre-performance test.
Results	Found that class drill in sight reading using either the vertical method or the ensemble approach were not significantly more effective than more traditional material. Found that the observed improvement in sight reading was significantly related to the number of previous college theory classes, and significantly related in the area of duration with the instrumentalists having the highest mean gain

Reliability	Shake Sight Reading Test inter-judge correlation = .867-.967 Reliability coefficient is .961 MAP reliabilities of composite scores for unselected subjects = .90 (for grade 4) - .96 (for grade 11); median of .94
Validity	-
Reference	Micheletti, L. H. (1980). *An assessment of the vertical method and the ensemble approach for teaching sight reading to secondary class piano students* (Doctoral dissertation, University of Miami). (UMI No. 8100741)

Title	Contributions of selected music skills to music sight-reading achievement and rehearsed reading achievement
Author	Ross Edwin Miller
Year	1988
Purpose	• To investigate the extent of seven selected music skills and seven personal characteristics related to eight indicators of music reading ability. Sub-problems: To what extent can 4 kinds of sight-reading ability (pitch, rhythm, expression, and all three combined) be predicted by the 7 selected music skills and how is the prediction affected when subjects are grouped according to selected characteristics including aptitude, attitude, musical background, field-dependence-independence, hand-eye coordination, and instrument family? To what extent can 4 kinds of rehearsed reading ability be predicted by the 7 selected music skills and how is the prediction affected when subjects are grouped according to selected characteristics including aptitude, attitude, musical background, field-dependence-independence, hand-eye coordination, and instrument family?
Measurement	• Grasp of tonality (MAT 2, part 2) • Pitch and rhythm audiation from notation (MAT 2, parts 3a and 3b) • Steady beat ability (Drake MAT) • Musical memory (Measure of Musical Abilities) • Musical background (Musical Background Test) • Attitude toward music (Music Attitude Inventory) • Field-dependence-independence (Group Embedded Figures Test) • Hand-Eye coordination (Frostig Movement Skills Test Battery) • Musical aptitude (Measures of Musical Abilities) Researcher and band director present for all tests, with researcher administering all individual tests. Tests were scored according to the printed test manuals, with exception of the WFPS. WFPS was scored to yield pitch, rhythm, and expression scores as well as a score in which only measure with no pitch, rhythm, or expression mistakes were counted (the "normal" score). 4 scores each for sight-reading and rehearsed reading were produced. Researcher designed and scored the Cognitive and Pattern Recognition tests Musical Background Test: Number of "yeses" were counted Music Attitude Inventory: 1-5 points for each answer (strongly disagree to strongly agree, respectively) Note Naming Subtest: Total number of correctly named notes from both clefs Rhythm Subtest: Number of correct measures with correctly placed bar lines Music Symbols and Terms Subtest: Number of terms and symbols correctly matched with definitions Fingering Subtest: Total number of correct fingerings Pattern Recognition Subtest: Number of notes which were contained in correctly identified patterns; 1 point for each note Rehearsed Reading Test: Scored in identical manner to sight-reading test (WFPS) with pitch, rhythm, expression and normal scores obtained For performance tests, researcher listened to each test at least 3 complete times (with

	difficult passages listened to many more times): once for pitch errors, then rhythm errors, then expression errors. A random sample of 12 sight-reading tests (10% of subjects) was scored by a second judge using same procedure, to check accuracy of scoring
Assessments Used	Music Achievement Test 2 (MAT 2), part 3a (Colwell, 1986) MAT 2, part 3b MAT 2, part 2 Drake Musical Aptitude Tests, form A (Drake, 1954) Measure of Musical Abilities, parts 2 and 4 (Bentley, 1966) Musical Background Test (Svenglais, 1978, unpublished) Music Attitude Inventory (Tomcala, 1977, unpublished) Group Embedded Figures Test (Oltman, Raskin, & Witkin, 1971) Frostig Movement Skills Test Battery, tests 1-3 (Orpet, 1972) Music Student Questionnaire (Author-composed) Watking-Farnum Performance Scale, form A (Watkins & Farnum, 1954) Cognitive Test (Author-composed), with 4 subtests: Note naming, Rhythm, Music symbols and Terms Matching, and Fingering
Description	
Subjects	123 students in sixth and seventh grade band (11-13 years of age). All subjects had just 2-4 years of school study.
Instruments Used	Wind instruments: 83 woodwind and 40 brass players
Stimuli	See Assessments Used Pattern Recognition Test: 3 songs: "America", "Yankee Doodle Boy", "This Old Man"; all found in elementary wind instrument method books Melody Pattern Subtest: "Up on the Housetop", "Bill Bailey"
Apparatus	*Realistic* electric condenser stereo microphone with *Teac* cassette deck for recording WFPS performances *SCSS* statistical software program (SPSS Inc., 1980) *Vax* computer system to run SCSS program *Philips* model #7941 receiver, *Teac CX310* stereo cassette deck, pair of *AR18* speakers for sound recordings during testing
Method	(Cognitive Test and Pattern Recognition Test were piloted with 45 subjects ahead of the school year) Entire testing program was 2 hours. The subjects were tested in 3 large sessions of 45-50 minutes each, followed by individual sessions of 20 minutes each for sight-reading, rehearsed reading, and hand-eye coordination. Larger group testing was done in the respective band rooms, with chairs and tables or desks for comfort with writing. Day 1: Note Naming Subtest, Rhythm Subtest, Drake Musical Aptitude Tests (form A), Music Symbols and Terms Subtest, Fingering Subtest Day 2: MAT 2 (parts 2 and 3), Rhythm and Melody Pattern Recognition Subtests Day 3: Bentley Measures of Musical Ability and Group Embedded Figures Test Also, 1 sixth grade group read Musical Background test silently on second day. The Music Attitude Inventory was completed at home. The seventh grade classes completed the Music Student Questionnaire, Musical Background test, and Music Attitude Inventory at home. Individual tests: Each subject took the sight-reading test, the rehearsed reading test, and the hand-eye coordination test. Hand-eye coordination subtests included the bead stringing subtest, the fist/edge/palm subtest, and block transfer subtest. The WFPS instructions were read to students. All subjects sight-read exercises 3, 5, 6, 7, 8. Immediately following, exercises 6 and 8 were practiced for 8 minutes and performed for rehearsed reading evaluation. Note-Naming Subtest: Subjects wrote letter names for notes (treble and bass) in spaces provided. There was a 3 min time limit Rhythm Subtest: Subjects were to divide 10 lines of rhythm patterns into complete measures according to the meter given. Subjects were given 4 min to complete

	Music Symbols and Terms Matching Subtest: Subjects matched the definition with the correct symbol or term Fingering Subtest: Subjects notated correct fingerings on 20 notes, written specifically for their instrument Pattern Recognition Test: Subjects circles patterns in the music and numbered them in similar groups. Time limit of 5 min
Results	Found that the general model of cognitive music skills-aural visual rhythmic skills-rhythm pattern perception skills is a logical combination of music skill, which could serve as basis for development of music reading instructional programs. Found that rhythmic aural-visual discrimination is more strongly related to instrumental sight-reading and rehearsed reading scores than melodic aural-visual discrimination. Rhythmic aural-visual skills show greater promise of affecting overall music reading skills with wind instrumentalists. Found that as music becomes more difficult, cognitive music skills appear to contribute increasingly to music reading achievement. Found consistent positive correlations among personal characteristics and music reading achievement. This means that information about personal characteristics can be of some use to the music teacher in prediction of music achievement
Reliability	Tests used had to have reported reliability of at least .70 (See Assessments Used)
Validity	-
Reference	Miller, R. E. (1988). *Contributions of selected music skills to music sight-reading achievement and rehearsed reading achievement* (Doctoral dissertation, University of Illinois at Urbana-Champaign, 1988). (UMI No. 8908779)

N

Title	'Musical brain' revealed by high-field (3 Tesla) functional MRI
Author	Tsutomu Nakada, Yukihiko Fujii, Kiyotaka Suzuki, Ingrid L. Kwee
Year	1998
Purpose	• To investigate the cortical areas subserving music literacy with high-field (3 Tesla) functional magnetic resonance imaging (fMRI)
Measurement	• To compare activation pattern associated specifically with music score reading with that associated with reading text in a subject's primary and secondary language • Pattern of cortical activation • Pre-Test: Musical literacy measured by subjects sight-reading randomly selected music scores on piano
Assessments Used	Edinburgh inventory for confirmation of right-handedness of subjects
Description	
Subjects	8 pianists (Ages 22-29). Primary language was Japanese; secondary language was English
Instruments Used	Piano (for pre-test of musical literacy)
Stimuli	Three conditions: Non-specific pictorial image (control state), presented text in either English (E) or Japanese (J), or music score (M)
Apparatus	*General Electric* (Waukesha, Wisconsin, USA) *Signa-3.0 T* system with *Advanced NMR (ANMR) EPI* module for all fMRIs *SPM96* (Wellcome Department of Cognitive Neurology) for analysis of images
Method	Each session was nine 3 0s epochs configured in the box car alternative sequences
Results	Found the activated areas common to all 3 reading modalities included auditory and visual association cortices. There was an area in the occipital cortex activated exclusively by reading music. The activated area common to all 8 pianists was identified as cortex adjacent to the right transverse occipital sulcus. Data supports notion that, as far as reading process is concerned, music shares a large portion of cortical processing with language. Music reading recruits additional cortical areas, especially those adjacent to right transverse occipital sulcus.
Reliability	-
Validity	-
Reference	Nakada, T., Fujii, Y., Suzuki, K., & Kwee, I.L. (1998). 'Musical brain' revealed by high-field (3 Tesla) functional MRI. *Neuroreport, 9*(17), 3853-3856.

P

Title	An experiment in the measurement of certain aspects of score reading ability
Author	Keith Areatus Pagan
Year	1970
Purpose	• To explore the possibility of devising a test procedure to measure the abilities of musicians to form accurate aural images of chords in a music score Can a group measure of the abilities of musicians to form accurate aural images of chords in a musical score be developed? Sub-problems: 1. If a test of chord-hearing ability is feasible, can it be validated against pertinent criteria? 2. If a test of chord-hearing ability can be developed which ahs concurrent validity, can it be demonstrated to be reliable to a reasonable degree? 3. If such a test if developed and is demonstrated to measure this skill with some degree of accuracy, then what is the relationship of this measure of ability to the subject's a. performing medium, b. piano training and experience, c. voice training and experience, d. formal training in music theory, e. total amount of music experience. 4. To what degree do changes in the characteristics of the stimulus relate to the measured ability? a. length of item, b. mode of item, c. change in the aural stimulus form the visual stimulus, d. type of change in aural stimulus
Measurement	• The ability to form an accurate mental image of the sound of the vertical aspect of a musical score (measuring instrument was limited to the measurement of this specific ability) Measuring instrument provides a means of measuring the non-keyboardist's ability to perceive the vertical aspect of the score. Teacher Rating Scale: Author-developed after fruitless search for reliable, validated measuring instrument to be used as an external criterion. 10-point ranking scale for use by teacher for rating the ability of each subject in harmonic dictation and harmonic aural analysis was devised. 1 point assigned to each correct answer and 0 assigned to each incorrect answer. Section B items had a possible score of 2 for each item. Pilot Study: Student Response Sheets scored by marking the errors and totaling the number of correct items, resulting in a raw score for each subject
Assessments Used	-
Description	
Subjects	Pilot Study: 12 graduate music students Form II: 79 subjects in basic musicianship courses at Indiana University and Pasadena College. These subjects had worked with a teacher or drill assistant over a period of time utilizing harmonic dictation, chord analysis, and other drills involving the vertical structure of music. *Main test: Sophomore music students (This was not described, so does not seem like there was a test after Form II.)*
Instruments Used	Baldwin Electronic Organ (Model 9), with the following stops: 8' Diapason, 8' Stopped Flute, 4' Flute. Organ was used because of pitch stability, sustaining power, and constant dynamic level for each pitch
Stimuli	Pilot study: 4 practice items for each section were composed. Section A – 60 items (30 items each appearing twice), 4-note chords, written on 2 staves with 2 notes to each staff, major and minor triads and inversions to be used. ½ of the items were changed in the aural stimulus. 2 types of changes to be used: a. same chord type, but different position or inversion, b. single alteration – one note altered by either a step or half-step in either direction. Only one type of change used in an item. 10 key signatures were used (no accidentals, 1 sharp – 4 sharps, 1 flat – 5 flats. Section B – 60 items (30 items each appearing twice), 4-note chords, 3 chords per item, rhythm factor is constant, 20 item pairs in major tonality with 10 item pairs in minor tonality. ½ items to contain alterations in aural stimulus. Only one chord per

	item to be changed, with first chord to be same in both aural and visual stimulus to be point of reference. Same key signatures and types of changes as Section A. Only diatonic triads, dominant sevenths, and inversions used. Phrases can begin on tonic, sub-dominant or dominant triad. ½ of changes occur in 2nd chord and ½ in 3rd chord. Aural stimuli were prerecorded. Form II: Same types of test items used as in Pilot Study; 30 items each for Section A and Section B.
Apparatus	Pilot Study: Acetate material for use in slides by a non-photographic process, with stimuli drawn with black india ink, *Philco Model TR 200* tape recorder for recording aural portion, *Argus 500* semiautomatic 35mm projector fitted with *Ibsor D.R.P.* camera for projecting stimuli (this was the tachistoscope), 40"x40" *Daylight* Projection Screen, and 70"x70" Projection Screen Form II: *Philco Model TR 200* tape recorder for aural stimuli, *Ampex Model 900* used for administration during the test, *Argus 500* semiautomatic 35mm projector with lens-mounted tachistocopic device, 40"x40" *Daylight* projection screen.
Method	Pilot Study: Subjects completed subject information sheet. Then turned to Response Sheet for Section A. The recorded instructions were played, and practice items were given. Subject had a chance to ask any questions, and then proceeded with Section A of the test. Subjects turned to Response Sheet for Section B. Instructions were played and practice items given. Section B of the test was given. Form II (of Pilot Study): Same testing procedure as Pilot Study, except for change that after verbal instructions were given and practice item projected, there were additional practice items with tachistoscopic projections, to help accustom subjects to seeing the 4-voice chord at faster speeds than used in the test. Experimental Test Procedure: 1. Number of the item is read, 2. Visual stimulus is presented, 3. Brief silence, 4. Aural stimulus is played, 5. Subject response
Results	Pilot Study: Percentage of correct responses was 67.2% for Section A and 68.5% for Section B. For Pilot Study and Form II, found similar results that keyboard study and experience had highest correlations with performance on the test. Voice study and vocal experience had lowest correlations. Found that the test was reliable and had appropriate validity. Author calculated a hypothesized Form III of the test, using some of the raw scores from Form II, to further improve the test
Reliability	Pilot Study: Section A = .925; Section B = .933 Form II: Index of reliability r = .90
Validity	Form II: Concurrent validity rho = .51
Reference	Pagan, K. A. (1970). An experiment in the measurement of certain aspects of score reading ability (Doctoral dissertation, Indiana University, 1970). *Dissertation Abstracts International*, A 31 (01), 1313.

Title	Relative effectiveness of two approaches to rhythm reading for fourth-grade students
Author	Mary Palmer
Year	1976
Purpose	• To study the effectiveness of two rhythm reading approaches
Measurement	• Pretest: Written achievement in rhythm reading and performance achievement in rhythm reading Each subject evaluated by a panel of 3 elementary music specialists, using a 5-point rating scale. The specialists evaluated independently, and scores were averaged
Assessments Used	Pretest: MAP (Gordon); Rhythm reading achievement measure: A combination of written achievement in rhythm reading, with "Meter Discrimination" section of MAT I (Colwell), "Auditory Discrimination in Music" from MAT II (Colwell), and "Rhythmic Concepts: Reading Recognition" from Iowa Test of Music Literacy, Level 1 (Gordon) Posttest: Rhythm reading achievement measures again
Description	
Subjects	Six classes of children in grade 4, from 3 elementary schools. Two classes are the control (38 students total), in one school. One class from each of the other schools is assigned the Richards method (48 students total), and one class from each of the other schools is assigned the Gordon method (50 students total)
Instruments Used	Unknown
Stimuli	Mary Helen Richards' rhythm reading method, based on the Kodaly system: *Threshold to Music* Edwin Gordon rhythm reading method
Apparatus	-
Method	For 5 months, experimental classes met for three 20-minute periods each week. Two experimental treatments were administered: 1. Mary Helen Richards' rhythm reading method, 2. Edwin Gordon's method. Control classes did not have a special program
Results	Significant difference between the control and experimental groups, with rhythm reading achievement. Found that written and performance achievement are two separate aspects of rhythm reading achievement
Reliability	-
Validity	-
Reference	Palmer, M. (1976). Relative effectiveness of two approaches to rhythm reading for fourth-grade students. *Journal of Research in Music Education, 24*(3), 110-118.

Title	The early development of sight-reading skills in adulthood: A study of eye movements
Author	Marjaana Penttinen and Erkki Huovinen
Year	2011
Purpose	• To elucidate the early stages of learning to read music in adulthood by examining the various measures of fixation time in elementary sight-reading tasks
Measurement	• Overall fixation time • Development of sight-reading skills through eye movements, for measures 2, 3, 4 only • Piano performance errors analyzed with sequencer files: all errors concerning one of the four critical notes were treated without distinguishing between pitch error types • Measured for temporal characteristics Three separate measurement sessions for each subject: beginning of training, after 2nd week of training, end of academic year Tested by experimenter
Assessments Used	Pretest: Written questionnaire for musical background information Posttest: All groups participated in skills test – ability to produce RH melodies, practiced from music notation

Description	
Subjects	49 BA education students in a compulsory music training period (30 included in final analysis). 15 were novices with no previous experience in playing instruments or able to read musical notation (target group). The other 15 had taken lessons for at least 1 year (control group)
Instruments Used	African drums, glockenspiel, piano
Stimuli	Twelve 5-measure author-composed melodies in C major. Each ended on E4 with range from C4 to G4. The fingering for first note was marked. All melodies were quarter notes with a final whole note. Movement primarily stepwise, with a perfect fourth and a perfect fifth
Apparatus	*Yamaha* electric piano for performances Computer screen for viewing *Power Tracks Pro Audio* for recording performances *Tobii 1750 Eye Tracker* for measuring eye movement Metronome for sight-reading measurement *ClearView 2.7.1* analysis software for analysis of eye movements
Method	Treatment: Administered in four 8-week long periods, with lessons on music didactics in small groups of approximately 16 subjects. During first three periods, the subjects had 20 hours of piano lessons in groups of 5-7. During fourth period, subjects had 16 hours of music lessons in groups of 15-20 to practice rhythms, melodies, and harmonies, and making simple arrangements with instruments Measurement Sessions: Sat at piano, eye tracker was calibrated, and subjects then sight-read 4 short melodies in time with metronome
Results	The novice group had skill development between first and third measurement. The novice group did not demonstrate sensitivity to metrical structure of melodies with consistent patterns of total fixation time, though this improved with practice. Intervallic skips in task melodies cause novices to increase their fixation times toward the end of the melody. This also improved
Reliability	-
Validity	-
Reference	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.

Title	Unexpected melodic events during music reading: Exploring the eye-movement approach
Author	Marjaana Penttinen, Erkki Huovinen, Anna-Kaisa Ylitalo
Year	2012
Study I: Pilot	
Purpose	• To test a research design suitable for examining the eye-movement indicators of unexpected melodic events
Measurement	• Eye-movement effect of unexpected musical events during a temporally constrained, simply music reading task • Fixation: Gaze dwelled within a 40 pixel radius for 60 ms + Only measured bars 2-7 as it was truly *prima vista* sight reading
Assessments Used	-
Description	
Subjects	5 female subjects (age 22-41) chosen from a group of 49 Finnish university students participating in a larger study on eye movements in music reading (Penttinen & Huovinen, 2009; 2011). These 5 subjects had extensive musical training and music reading ability, successful eye-movement and performance recordings for the 3 melodies used, and temporal and melodic accuracy. All had formal instruction for at least 10 years, all had taken piano lessons
Instruments Used	Piano

Stimuli	"Mary Had a Little Lamb", in key of C Major, with variations of the notated melody. All written for RH (C4-G4), with one finger for each note. The variations did not present salient visual complexities nor alter the original rhythm
Apparatus	*Tobii 1750 Eye Tracker* (Tobii Technology AB; Stockholm, Sweden) *Finale* music notation software Computer for viewing music *Yamaha* electric piano *Power Tracks Pro Audio* for recording performances *ClearView 2.7.1* analysis software
Method	Each subject participated in 3 measurement sessions over 9 months. The subjects participated individually, with an experimenter present. The subjects were introduced to the equipment and were seated at the piano. They first played short diatonic melodies in C Major; quarter notes performed at 60 M.M. They then played the original melody of "Mary Had a Little Lamb" to a metronome at the same speed. They were asked if they recognized the melody (they did). They were instructed to play the next piece after 4 metronome beats. In each measurement they played one variation.
Results	Did not find that the subjects spent more time on altered bars in all 3 variations. However, the altered bar and preceding bar in each variation together collected nearly 50% of all incoming saccades (movement from last measure into next measure). The unexpected events in the melody did cause visual processing to re-adjust.
Reliability	-
Validity	-
Study II	
Purpose	• To statistically examine how unexpected melodic deviations affect the allocation of fixation time and saccadic eye movements for musically experienced subjects
Measurement	• Fixations of the staff system • Relative fixation times • Incoming saccades
Assessments Used	-
Description	
Subjects	34 (originally 40, but data reported only for 34) education majors minoring in music education at a Finnish university (n=21) and music students of performance at a Finnish arts academy or conservatory (n=13); ages 19-37 years old; 22 were female
Instruments Used	Piano
Stimuli	The original version of "Mary Had a Little Lamb", as well as variations 1 and 2 from Study I
Apparatus	*Sibelius 6.2.0* music notation software 23" widescreen *TFT* monitor to present melody *Tobii TX300 Eye Tracker* (To ii Technology AB; Stockholm, Sweden) *Yamaha* electric piano *Power Tracks Pro Audio* sequencer software for recording
Method	Subjects tested individually with experimenter present. First, subjects were introduced to the laboratory setting. They played the original melody first, to a metronome set at 60 M.M., using only RH. Then they practiced by the presentation of a series of written instructions followed by short, simple melodies in C Major (author-composed). Before each melody, the subject looked at a cross, which marked the location of the first note two second in advance of the appearance of the staff. After the staff appeared, the subject waited two beats before performing. For the actual test, the subject was told that they would perform four versions of "Mary Had a Little Lamb" at the same tempo, and that some of the melodies would have slight alterations from the original melody. The order of the melodies was changed for two conditions.
Results	Found that alterations of the melody affected the relative fixation time and the distribution of incoming saccades. The effects were observed for the bars around the altered bars. When the melody was altered in the middle of the melody, the relative

	fixation time was decreased to the bar after. The altered bars increased the proportion of incoming saccades to the bars before them. Immediately following the altered bar, there was a decrease in the relative amount of incoming saccades.
Reliability	-
Validity	-
Reference	Penttinen, M., Huovinen, E., & Ylitalo, A. (2012). Unexpected melodic events during music reading: Exploring the eye-movment approach. In E. Cambouropoulos, C. Tsougras, P. Mavromatis, K. Pastiadis (Eds.), *12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music* (pp. 792-798). Retrieved from http://icmpc-escom2012.web.auth.gr/sites/default/files/papers/792_Proc.pdf.

Title	Reading ahead: Adult music students' eye movements in temporally controlled performances of a children's song
Author	Marjaana Penttinen, Erkki Huovinen, Anna-Kaisa Ylitalo
Year	2015
Purpose	• To study the potential effects of musical background and selected local music-structural details on visual processing of simple score in effortless and fluent, audibly similar performances. • To explore gaze activity (the number of inspected visual beats between two beat onsets) Is gaze activity affected by musical background, melodic alterations, or simple rhythmic patterns in temporally controlled performances? • To examine if the eye-hand span is affected by musical background, melodic alterations, or simple rhythmic patterns during accurate performances in controlled tempo
Measurement	• Performance errors: Note substitution, note addition, late note (with respect to metronome) • Fixation • Eye-hand span and gaze activity: Assigned fixations to particular note symbols. Visual field of measures 1-8 divided into 12 Areas of Interest (AOIs), each corresponding to quarter-note beat
Assessments Used	-
Description	
Subjects	38 (originally 40, but data reported only for 38) education majors minoring in music education at a Finnish university (n=24) and music students of performance at a Finnish arts academy or conservatory (n=14); ages 17-37 years old; 25 were female
Instruments Used	Piano
Stimuli	The original version of "Mary Had a Little Lamb", notated in C major, as well as 2 slightly altered versions, where target bar did not greatly deviate in visual appearance from original
Apparatus	*Sibelius 6.2.0* music notation software 23" widescreen *TFT* monitor to present melody *Tobii TX300 Eye Tracker* (To ii Technology AB; Stockholm, Sweden), both eyes tracked at 300 Hz *Yamaha* electric piano *Power Tracks Pro Audio* sequencer software for recording
Method	Subjects completed questionnaire concerning musical background. Began with (i) familiarization phase where subject was presented with original version of melody on computer screen and asked to practice it with metronome set at 60 bpm, using RH. The (ii) practice phase had the subject presented with written instructions and short, simple melodies on computer screen, while recording eye movements. Before each melody, subject looked at cross, and marked the location of first note 2 s before appearance of staff. When staff appeared, subject waited 2 s (beats of metronome)

	before playing. The (iii) test phase had subject performing 4 versions of the original melody in same tempo, but that there would be slight alterations. Subject was to play all stimuli. After experiment, subject saw video recording of their eye movements and discussed music-reading behavior with experimenter. Session lasted 25-35 minutes
Results	Found that average EHS length is 1 s (consistent with other studies), even with extremely simple sight-reading tasks. Found expertise effect of shorter average fixation durations, which may also be true in temporally controlled conditions. Expertise effect of frequent use of extended spans and greater gaze activity. Found that with both groups, gaze activity increased during eighth-note patterns; therefore, denser patterns measure as a local increase of gaze activity. With 1-measure melodies that were altered, there was a tendency towards diminished gaze activity and shorter EHS. So, expertise effects of eye movements may be 'locally evened out' by just a slight violation in melodic expectations.
Reliability	-
Validity	-
Reference	Penttinen, M., Huovinen, E., & Ylitalo, A. (2015). Reading ahead: Adult music students' eye movements in temporally controlled performances of a children's song. *International Journal of Music Education, 33(*1), 36-50.

Title	Processing of local and global musical information by unilateral brain-damaged patients
Author	Isabelle Peretz
Year	1990
Purpose	• To investigate melody processing in unilaterally brain-damaged patients. • To assess the idea supported by data fro normal that the left hemisphere is better equipped for dealing with local features of melody and the right hemisphere for arriving at global melody representations. Addressed 2 questions: Given that music perception can be affected by a unilateral vascular lesion (Benton, 1977), what can these impairments tell us about the contribution of each hemisphere to music processing? What can these deficits tell us about existing models of normal music processing?
Measurement	• Melody processing, by manipulating the availability of contour and metre
Assessments Used	-
Description	
Subjects	20 unilateral brain-damaged patients and 20 normal controls (NC). All brain-damaged patients had suffered from a first unilateral cerebrovascular lesion. Controls were matched as closely as possible in age, sex, education, musical background and occupation. None of the subjects were musicians.
Instruments Used	-
Stimuli	Set of 6 subtests constructed from same pool of 24 novel musical sequences. They were tonally structured and made up of 2 phrases following Schoenberg's (1942) principles. Half in duple metre, half in triple metre. The 2-phrase sequences used in metric condition; all other conditions only used second phrase of each sequence. The second phrases were 4 measures, lasted 4 s, and had 8-19 tones. Contour-violated, contour-preserved, and transposed melodies in 3 different conditions – this manipulation applied to second phrases. One manipulation was transposing each sequence to a near key. Another manipulation was a creation of a contour-violated alternate melody by modifying pitch of one tone so that it changed the pitch direction of surrounding intervals, while keeping in the original key. Third manipulation consisted of creating contour-preserved alternate melody of these contour-violated melodies by modifying same critical pitch to same extent, but in keeping with the original contour and scale. 3 sets, each consisting of 2 practice trials and 24 experimental trials. For rhythmic task, the stimuli were derived from second phrase of the 24 sequences

	used in metric task. 2 manipulations were applied to these phrases: One consisted of removing all pitch variations by tapping the temporal patterns; care was taken as to mark the metre by placing an intensity accent on first beat of each measure. These patterns served as targets. Second manipulation consisted of modifying half of these targets by interchanging the durational values of 2-3 sounds, while keeping metre and total number of sounds identical. Set of 2 practice and 24 experimental trials constructed with these temporal patterns. For recognition tests, from initial set of 24 melodies, 10 were selected randomly. Each
Apparatus	Piano and *Revox A 77* tape recorder for playing and recording of stimuli Stimuli was presented on same tape recorder during experiment Speaker in front of subject for prerecorded tapes
Method	Each trial had a warning signal and target melody followed, after 2 s silent interval, by a comparison melody. Duration of inter-trial interval was 5 s. For each set, there were twice as many 'different' as there were 'same' trials. For rhythmic trials, half the of the trials had the target pattern repeated after a 2 s silent interval, and on the other half it was followed by its different comparison version. Subjects were tested individually in single session. They listened to prerecorded tapes. Session began with musical 'entrance' of 10 excerpts taken from commercial recordings. They were to judge whether or not they were familiar with music presented. Then they were presented either with the pitch organization conditions or with the temporal tasks. For the former, all subjects were presented first with contour-violated condition, and then contour-preserved condition, followed by transposed condition. For latter, order of presentation was counterbalanced for temporal tasks across subjects in each group. Session ended with recognition test. For all conditions (except metric task and recognition test), subjects required to perform 'same-different' classification task. They had to judge whether target sequence and comparison sequence were the same or not. For transposed condition, subjects were instructed to disregard the fact that the melodies were played at different pitch levels. Before each condition, 2 practice trials were presented and a symbol (tree) was used on the top of response sheet to help in choosing right answer. Practice trials could be repeated until subject understood what was required. Then they completed the 24 experimental trials. Subjects responded by marking crosses on the response sheet below corresponding symbol. For metric task, subjects told they would hear waltzes and marches and had to discriminate them along that dimension. They had to mark a cross below the corresponding symbol. They were encouraged to tap along with what they perceived to be underlying beat of each sequence. For recognition test, subjects instructed that they would be hearing some of the tunes they had been listening to during the session, mixed with new ones. They were to respond 'yes' if they recognized the tune and 'no' otherwise
Results	Found that present findings highly consistent with the view that left hemisphere deals with local features of melody and right hemisphere deals with global melody representations, and further qualify the functional differences. Found that NC and LHD used contour cue, while RHD found to exhibit little sensitivity to presence of contour as a discrimination cue. This indicates that a lesion in the RH interferes with the possibility of representing and/or retrieving the contour of melodies. Found that a lesion, whether in the R or L hemisphere, can impair local processing of melodies. Found the pattern of association and dissociation suggests that the processing of local intervals is dependent on intact contour processing. For temporal dimension, only the discrimination of durational values (rhythm) was found to be impaired by lesion in either hemisphere, which spared metric interpretation of the sequences. Evidence of double dissociation between the processing of the pitch dimension and

	processing of rhythm, providing further support for need to fractionate musical perceptual abilities. Main conclusion is that cooperation of the 2 hemispheres is required to achieve efficient organization of pitch sequences, while admitting that R lesion is more disruptive to this function than L lesion
Reliability	-
Validity	-
Reference	Peretz, I. (1990). Processing of local and global musical information by unilateral brain-damaged patients. *Brain, 113*, 1185-1205.

Title	Learning to read music: Children's use of structure in pitch notation
Author	Anne D. Pick, Marsha G. Unze, Steven Metz, and Rose Mae Richardson
Year	1982
Experiment 1	
Purpose	• To assess whether the systematic relations of the pitch notation system to the keyboard spatial arrangement, and of the pitch notation system to the notes of the melody affect children's initial learning of pitch notation
Measurement	• Two scores: Number of repetitions needed to achieve one 'perfect' playing of the melody, and number of repetitions to reach same criterion in transfer phase
Assessments Used	-
Description	
Subjects	60 children from grade 1 and 2 classrooms
Instruments Used	A set of flat rectangular metal bells, similar to a glockenspiel, with pitches C, D, E, G; hit with a hammer
Stimuli	Notation printed on a 5-line staff, with small triangle used for pause or rest. The focus was on pitch notation. Notation was different than Western notation: C on lowest space, D on second line, E on second space, G on third space. Each of the 6 melodies was 12-14 notes: 2 familiar, 2 unfamiliar, 2 based on familiar tunes with alterations
Apparatus	-
Method	Normal Condition: Bells arranged linearly in ascending order. Rearranged Condition: Bells disordered. Single Pitch Condition: All bells were C. Total of 6 subgroups with 3 experimental conditions, and familiar or unfamiliar training song in each. 10 children in each subgroup. Experimenter and child sat at a table with a bell box and music stand Child learned how to strike bell. Experimenter then played their particular training piece for them. Initially, the correspondence of notes to playing was explained generally, then experimenter pointed to each note on staff and asked child to play the corresponding bell. Then they played through again with the experimenter pointing to the staff and only prompting child with correct bell choice when necessary, or correcting any errors. One group then had experimenter help continually as needed, while the other group did not. Procedure continued until child played the song through once without error nor prompts
Results	Evidence that children do perceive structure in notation. With pitch notation and 'keyboard' distorted (rearranged condition), more training was required. The single pitch condition learned the task quickly but it did not help them with the second, transfer song. There was a minimal transfer advantage for the children with a familiar training song. Followed by a supplementary study using normal or rearranged conditions, where notes were coded by four colored circles and associated with a particular bell
Reliability	-
Validity	-
Experiment 2	
Purpose	• To assess how children perceive the relation of notation to keyboard and pitch changes

Measurement	• The differences of ease of learning and degree of transfer between the two stimuli formats: one note versus score Each child received a score for number of repetitions of the training melody, a score for a number of repetitions of transfer melody, and a score for the difference between training and transfer trials
Assessments Used	-
Description	
Subjects	44 children in grades one and two
Instruments Used	Bells with pictures D, E, F, G, A
Stimuli	The notation was presented to either with only one note visible at a time (Single Note), or available for scanning (Standard). Stimuli included the first 13 notes of “London Bridge” and “This Old Man”. Each was the training melody for half of the participants, and the transfer melody for the other half
Apparatus	Box used for stimuli condition where one note was presented at a time. The experimenter turned the paper scroll to reveal the next note
Method	The Normal Condition had information about structure and pitch notation manipulated. For conditions with two training notation formats: Standard or Single Note. The children had to achieve two perfect renditions of their training melody before proceeding to the transfer piece, which also needed two perfect renderings
Results	Single Note Condition did not enhance or retired the children’s learning of the training melody. Could provide evidence that the perceived structure is not immediately helpful for learning the specific note-bell correspondences. Learning just a few notes and its relation to the keyboard provides predictability. It can be predicted that children will perceive sooner the equivalences that are systematic
Reliability	-
Validity	-
Reference	Pick, A. D., Unze, M. G., Metz, S., & Richardson, R. M. (1982). Learning to read music: Children’s use of structure and pitch notation. *Psychomusicology, 2*(2), 3346.

Title	Employing cognitive chunking techniques to enhance sight-reading performance of undergraduate group-piano students
Author	Pamela D. Pike and Rebecca Carter
Year	2010
Purpose	• To discover whether exposing subjects to different modalities of rhythm or pitch chunking techniques would result in improved sight-reading performance when compared with subjects who did not engage in chunking-drill treatments
Measurement	• Pretest and Posttest: Sight-reading skills Researchers observed through video recording and in-person, as well as observation of preparatory behavior 2 independent reliability observers analyzed 20% of pretests and posttests. Scoring done with modified WFPS (Watkins & Farnum, 1962), scoring RH rhythmic accuracy, LH rhythmic accuracy, RH pitch accuracy, LH pitch accuracy, and continuity. Subjects could earn 5 points for each beat of music
Assessments Used	Pretest and Posttest: Three 8-measure sight reading examples, with the posttest being more challenging Posttest: Attitudinal questionnaire
Description	
Subjects	43 music students in 6 intact piano classes at a university. Each class was randomly assigned to one of 3 study groups: Control group (A, n=15), Experimental group (B, n=13) and Experimental group (C, n=15)
Instruments Used	Piano
Stimuli	Pretest: 3 examples were taken from *Alfred’s Group Piano for Adults, Book 1, Teacher’s Edition* (2nd ed.) (Lancaster & Renfrow, 2004b). These were material that first-semester music majors would be expected to perform comfortable. C, G, D

	Majors, 3/4, 4/4, 2/4 meters (for pretest); 4/4, 3/4 meters (for posttest). Treatment: 12 examples. 2 were author-composed; 10 from textbook being used in class
Apparatus	*Finale* for rhythm and pitch drill notation *Yamaha Disklavier* piano for performances *MIDI* for recording performances and for accompaniment during final group performance Projectors to present sight reading
Method	Pretest and posttest: Subjects had 30 seconds to study each example silently before performing. After nine weeks of piano instruction, the three-week treatment began where are the primary investigator instructed each group in sight reading activities for the final 10 minutes of class, two times each week (60 minutes total). Each group sight-read 12 examples during the instructional period. Control group: No prior rehearsal of practice drills and without explicit reference to rhythm or pitch patterns. Experimental groups: Rehearsed rhythm or pitch drills prior to practicing the examples. Subjects practiced two examples during each class period. Group A: Instructed to play hands separately several times, then together. Group B: First performed a series of 1 - 4 rhythmic tapping drills per example, tapped twice. These rhythms were found in the example. That display showed the rhythmic patterns highlighted and circled. Group C: Same procedure as a group B but with pitch patterns. They were encouraged to see accompaniment patterns as chords, and melodic lines outlining chords were also highlighted
Results	No significant difference between the three groups for overall sight-reading scores. There were significant interactions shown between pretest and posttest scores for all groups. Concluded that the ability to perform rhythm patterns correctly was related to the ability to execute the pitches for these particular subjects who were exposed to chunking drills
Reliability	Inter-rater Reliability: Pretest = .98, Posttest = .99
Validity	-
Reference	Pike, P. D., & Carter, R. (2010). Employing cognitive chunking techniques to enhance sight-reading performance of undergraduate group-piano students. *International Journal of Music Education, 28*(3), 231-246.

S

Title	The identification and assessment of cognitive variables associated with reading of advanced music at the piano
Author	Doris Louise Salis
Year	1977
Purpose	Focus of the study is on some cognitive aspects of various senses used in sight reading: Visual sense (primary focus), aural sense (from cognitive perspective), visual perception process (with recognition that reading music at the piano also involves the kinesthetic sense of musician) Main purposes of study are: • To identify and assess particular cognitive variables which are correlates of advanced music reading at the piano • To compare performance on these variables in 2 groups of music readers: 'Good' and 'poor'
Phase I: Music Reading Test (MRT)	
Measurement	Each excerpt performed was recorded and the experimenter scored the medium and hard selections; 30 excerpts for each subject • Errors were categorized into tonal and rhythmic errors, with subcategories for each. Good readers scored above the median, and poor readers scored below the median • Scored testing time • Erors per minute
Assessments Used	-
Description	
Subjects	26 (13 male, 13 female; Average age = 27.15 years). All subjects either college keyboard majors, professional musicians, four paid professional piano accompanists, 7 who accompanied in their work without pay
Instruments Used	Piano
Stimuli	45 excerpts from standard piano repertoire; 5 styles of music. 9 excerpts for each composer: Bach, Mozart, Beethoven, Brahms, Impressionistic (Debussy and Ravel). For each excerpt, 3 professional musicians (including researcher) judged excerpts as easy, medium, or hard for reading level. Excerpts an average of 7 measures. All keys represented in the group of excerpts.
Apparatus	Grand Piano Binder with music Tape recorder for performances Stop watch for performances
Method	Subject at piano, with binder of music in front of them, and 30 s to study the music. The subject had to tell the experimenter if they had previously studied any of the music. The subject could choose their own tempo, but should note the tempo markings written on the music. Subjects did not have to play additional embellishments. They played the various excerpts while being recorded.
Results	Poor readers made more errors and played slower. They made both tonal and rhythmic errors and took more time to make the errors. Opposite is true of good readers. Not a highly significant difference between good and poor groups with errors per minute.
Reliability	Spearman-Brown split-halves method = .98
Validity	-
Phase II	
Measurement	Comparison of performances of each group with selected variables • Musical aptitude (musical memory and rhythm) • IQ (scores for verbal ability and for performance ability; subtests for digit span, block design, picture arrangement)

	• Word reading ability (combined score of vocabulary, comprehension, and labeled word reading skills, reading rate) • Mental imagery ability (ability to imagine in three dimensions) Correlated performance on variables with performance on MRT
Assessments Used	Musical aptitude: Drake Musical Aptitude Tests (Drake, 1954) Intelligence: Wechsler Adult Intelligence Scales (WAIS) Word Reading Ability: Nelson-Denny Reading Test Mental Imagery: Block Visualization Test (Aptitudes Research Project, University of Southern California, under U.S. Government Grant I.E. 1971)
Description	
Subjects	Same as Phase I
Instruments Used	-
Stimuli	Stimuli of various assessments used
Apparatus	Assessments and writing utensils
Method	Group or individual testing as described for each assessment
Results	Music reading test correlated most with Word Reading, IQ, and Musical Memory, though they are moderate to low correlations. Musical memory seems to share more with word reading skill and general IQ than with music reading directly. Block visualization represents a small share of the skills needed in actual music reading.
Reliability	May have more details per each assessment
Validity	May have more details per each assessment
Phase III	
Measurement	• Laterality effects during unilateral viewing of musical chords, and random dot patterns. Laterality tests correlated with scores on variables assessed in Phases I and II • Dots: Scores for left and right fields of total number wrong, and difference scores (difference between subjects' response and actual number of dots) • Chords: Two scores for each level of chunking and each complexity type in each field: Total note errors and totally correct chords
Assessments Used	-
Description	
Subjects	26 subjects (13 good and 13 poor readers) from Phase I
Instruments Used	Piano (for experiment II)
Stimuli	Dots: Random dot stimuli were drawn on cards, with 64 pairs of cards. There were 8 pairs each of 8 chunking sizes (3-10 dots) Chords: 63 pairs on cards. Three types of chords: TP (plain tonal chord without accidentals), TA (tonal with accidentals), AT (atonal with mixed accidentals and odd interval combinations)
Apparatus	Tachistoscope *Wurlitzer* electric piano (for chord experiment)
Method	Dots: Subjects centered their eyes on the field cross. When experimenter said "ready", they were to count the number of dots appearing randomly to the left or right of cross. They responded verbally. Chords: Subjects focused on centre cross, then experimenter said "ready", and subject played as many notes as possible, with fingers kept on keys, until experimenter said "OK".
Results	Dots: General preference for left field by both groups and sexes. Males significantly more accurate on dot patterns with 6-10 dots. Chords: General preference for right field. Good reading group did better. Both groups did better with run 2 on chords. There is a practice effect for poor and good readers. TP chords were better perceived.
Reliability	-
Validity	-
General Results	Best predictors for music reading ability are: IQ, word reading ability, note errors on

	Phase III chord test. Can assume that the cognitive ability represented by the chord test is not shared with IQ or word reading ability. Music readers' ability to 'chunk' stimuli seems to be specific to musical stimuli. Appears that good readers have developed ability to chunk musical stimuli and relate those chunks to a higher degree than poor readers.
Reference	Salis, D. L. (1977). *The identification and assessment of cognitive variables associated with reading of advanced music at the piano* (Doctoral dissertation, University of Pittsburgh, 1977). Retrieved from ProQuest Dissertations and Theses. (UMI No. 789605)
See Also	Salis, D. L. (1980). Laterality effects with visual perception of musical chords and dot patterns. *Perception & Psychophysics, 28*(4), 284-292.

Title	A comparison of prompts to aid rhythmic sight-reading of string students
Author	Rita S. Salzberg and Cecilia Chu Wang
Year	1989
Purpose	• To investigate the effectiveness of three prompts on rhythmic sight-reading ability
Measurement	• Prompts with rhythmic sight reading To score, each quarter note beat of measure was given value of 1 point. All exercises were 8 measure of 4/4 meter, so a total possible score of 32 points. Each performance was evaluated separately by one researcher and a second judge
Assessments Used	-
Description	
Subjects	46 string students; ages 8-16; range of 1-7 years of instruction on their instruments
Instruments Used	Violin, Viola, Cello, Double-Bass
Stimuli	Rhythmic sight-reading exercises at 10 levels of difficulty: Level 1 – quarter, half, whole notes, quarter rests. Level 2 – quarter, half, whole, dotted half notes, quarter, half, whole rests. Level 3 – eighth, quarter, half, whole notes, quarter, half, whole rests. Level 4 – 9 not described. Level 10 included all listed below, and sixteenth and dotted eighth notes. 4 exercises were written for each level; total of 40 exercises. Each exercise was 8 measures and played on open D string.
Apparatus	Metronome (60 pulses per minute) to prep each exercise Tape Recorder for performances
Method	Pretest for all students to place each at the proper difficulty level. They were grouped according to their years of instruction and played lower (if too difficult) or higher, until each reached their maximum difficulty level. 4 exercises at each level were given in sequence to 4 different subjects. The pre-test allowed examiners to confirm that each of the 4 exercises in a specific level was of comparable difficulty, and that the levels were an accurate progression of difficulty. From pre-test results, subjects were placed into 1 of 5 levels of rhythmic-reading ability. Subjects practiced each of the prompts in training sessions. There were 4 treatment conditions, which were prompts/aids to rhythmic sight-reading: 1. Foot tapping – Subjects tapped foot to quarter note beat, 2. Counting out loud – Subjects counted beats and eighth note subdivisions out loud (except for complex rhythms of group 5, where they could use any method they were comfortable with), 3. Tapping and counting simultaneously – A combination of the two prompts above, 4. No aid/prompt. Each subject tested individually, playing the 4 exercises of that particular level, with each exercise using a different prompt. The 4 treatments were randomly ordered. If the subject stopped, they continued at the same measure where they stopped. A metronome began each exercise and was stopped as soon as the subject started playing
Results	Found that only counting prompt was effective in rhythmic sight-reading, and improved sight-reading with 3 lower levels of difficulty only. Found that the counting

	condition was more effective for younger, less musically experienced subjects. At higher levels there was no difference among prompt conditions
Reliability	Reliability between the two scorers = .75-.97, with a mean of .885
Validity	-
Reference	Salzberg, R.S., & Wang, C.C. (1989). A comparison of prompts to aid rhythmic sight-reading of string students. *Psychology of Music, 17*(2), 123-131.

Title	Processing pitch and duration in music reading: a RT-ERP study
Author	Daniele Schön and Mireille Besson
Year	2002
Purpose	• To examine processing of pitch and duration in music reading, using electrophysiological and behavioral methods, specifically in the visual form • To determine whether pitch and duration are processed independently or jointly
Measurement	• ERPs • Electrooculography (EOG): Both horizontal and vertical
Assessments Used	-
Description	
Subjects	18 musicians, only 15 retained. All RH.
Instruments Used	-
Stimuli	Clef with key and time signature first as probe, presented for 500 ms. Blank screen for 1000 ms then target note presented on staff for 500 ms. 4 experimental conditions: 1. Target note congruous with both tonality and meter (PcDc), 2. Target note congruous with tonality but not meter (PcDi), 3. Target note congruous with meter but not tonality (PiDc), target note not congruous with etiehr (PiDi). 60 trials in each condition
Apparatus	Faraday box for subjects to sit *SVGA* color computer screen for presentation
Method	Subjects were to decide, as quickly as possible, whether the target note matched or mismatched with the tonal or metrical information in the probe. For Block 1, the accidentals or the numbers of the probe were colored green or red, and subjects were to judge whether target note matched/mismatched the colored dimension. In pitch task, they were asked if target note was most representative of key/tonality. In duration task, they judged whether note properly fitted duration of the measure as indicated by metric information. For Block 2, there was no color on the probe. The target note was red or green and subjects were asked to perform the pitch task or duration task, depending on the color given. For Block 2, the subjects therefore had to process both pitch and duration information.
Results	Block 1 was easier than Block 2. RTs were shorter in the pitch task than in the duration task, and for congruous than incongruous target notes. ERPs reflected congruity by a negative component to incongruous targets. The congruency of the target note on the irrelevant dimension did not have any effect on ERPs in either block. For ERPs, there may be an increased positivity in the duration task or an increased negativity in the pitch task (can consider either way, but researchers in this experiment consider the latter). Because the ERPs did not show any influence of the irrelevant dimension on the relevant dimension, it seems to favor the view that pitch and duration are processed independently
Reliability	-
Validity	-
Reference	Schön, D., & Besson, M. (2002). Processing pitch and duration in music reading: a RT-ERP study. *Neuropsychologia, 40*, 868-878.

Title	Visually induced auditory expectancy in music reading: A behavioral and electrophysiological study
Author	Daniele Schön and Mireille Besson
Year	2005
Purpose	• To investigate the processes involved in reading musical notation and to study the relationship between written music and its auditory representation. Goal to determine whether musicians are able to develop expectancies for specific tonal or atonal auditory events based on visual score alone. Wanted to determine whether modulations in amplitude of a late positive component (P600), as described in previous studies, are linked to a general mismatch detection process or to specific musical expectancies
Measurement	• Behavior (reaction times, or RTs, and error rates) • Electrophysiological data (event-related brain potentials, or ERPs) • Electrooculography (EOG): Both horizontal and vertical
Assessments Used	-
Description	
Subjects	18 musicians, but only 14 retained. All RH
Instruments Used	-
Stimuli	320 pairs of visual/auditory stimuli. Visual Stimuli: All quarter notes. Two types: 5 notes with a stable last written note, and 5 notes with an unstable last written note (with respect to tonal context). Keys used: C, F, Bflat, Eflat, G, D, A, B, d, f, fsharp, g, a, b). Last note was always an in-key note for stable stimuli Auditory stimuli: Three types: Final note either matched, or mismatched with last note of visual. Mismatches were either plausible or implausible Total of 6 experimental conditions: Stable or unstable notes at end with three types of auditory sequences: 1. Ending with notes that match visual sequence, 2. Mismatch but are musically plausible, 3. Mismatch and musically implausible. Specifically, in the visual-auditory mismatching condition: Either 2 and 5 was mismatched (plausible), or 3 and 6 was mismatched (implausible), with respect to the tonal context
Apparatus	Computer screen MIDI (*Korg, X5DR*) for auditory stimuli generation Faraday box for subjects to sit Computer to play auditory stimulus Scalp electrodes for EEG
Method	Subjects judged whether the last note of a 5-note auditory musical sequence matched/mismatched a score presented to them. They were seated in a Faraday box. They read score while computer played the notes. They were to judge as quickly as possible, whether the final note was a match or mismatch
Results	Accuracy and RTs had consistent pattern of results: Matching conditions were faster and with fewer errors. RTs were shorter for stable endings. With ERPs, mismatching conditions differed from matching conditions when visual stimuli had stable and unstable endings. With auditory-visual matching, the amplitude of early and late negative components (N2 and N5), and amplitude of positive component (P3), is larger in the unstable than stable ending. This shows that unstable matching events seem to be processed differently from stable matching events. With auditory-visual mismatching, there are larger N1, P3, and LPCs with implausible than plausible. Musicians are able to anticipate stable endings based on visual information, and able to do the same for unstable endings, though to a smaller extent. There are strong interactions existing between visual and auditory musical codes: The visual stimuli influences the way auditory musical sequences are perceived. "Although there is an early effect of visual information on auditory processing, the effect of syntactic incongruity is only reduced and does not disappear" (p. 702)
Reliability	-

Validity	-
Reference	Schön, D., & Besson, M. (2005). Visually induced auditory expectancy in music reading: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience 17*(4), 694-705.

Title	An fMRI study of music sight-reading
Author	Daniele Schön, Jean Luc Anton, Muriel Roth, Mireille Besson
Year	2002
Purpose	• To investigate the brain areas involved in music reading, using fMRI • To address the question of whether the processes involved in music sight-reading are independent from those at play when reading words and numbers • To evaluate the specificity of the brain areas, reading music notation was compared to reading verbal and number notations
Measurement	• Brain areas and specificity of these areas with different notations
Assessments Used	-
Description	
Subjects	9 right-handed subjects; 4 women, 5 men; ages 24-50 years. All had minimum 12 years piano playing experience
Instruments Used	
Stimuli	Condition 1: music notation with 5 notes, from do to sol, with thumb corresponding to do. Condition 2: verbal notation with 5 notes, from do to sol. Condition 3: Arabic number notation with 5 numbers, from 1 to 5. Stimuli in control conditions visually matched those in the experimental conditions: a quarter-pause on the musical staff, a short word (which was 'chut', meaning 'silence') and a zero. Each block was 26 stimuli; with 10 blocks for each condition, over 3 scanning sessions
Apparatus	*3 T whole-body imager MEDSPEC 30/80 ADVANCE* (Bruker) for imaging of brain *SPM99* statistical parametric mapping software for image processing and analysis *Montreal Neurological Institute* template for anatomical references and the realigned functional images of all subjects
Method	There were 3 experimental conditions and 3 control conditions. For all conditions, stimuli were presented visually one at a time, in a pseudo-random order, for 800 ms, with inter-stimulus interval of 150 ms. In experimental conditions, the RH played on a 5-key keyboard. In control conditions, subjects pressed a button with fourth finger each time a stimulus appeared. At beginning of each block, instructions were displayed to indicate type of notation to be presented (this is the only description of the methods)
Results	Found that overall, a similar pattern of brain areas is activated by the 3 notational systems: the parietal lobes bilaterally, the sensorimotor cortex of the left hemisphere, and the right cerebellum. For music versus words and music versus numbers, there were 2 major foci of activation: the right superior parietal lobule (SPL) and the intraparietal sulcus (IPS), mesial to the supramarginal gyrus. Another minor focus of activation for both contrasts was also found in the right visual cortex, close to the occipito-temporal junction. Therefore, they confirmed that when playing from music notation, well-defined right parietal regions are more involved than when transcoding from verbal or number notations. There are different ways of reading music and these can eventually be at work at the same time, but authors do not yet know whether the same cerebral parietal networks would also be involved when reading musical notation in order to sing or name notes. Found that the right occipital cortex seems to be important to music reading
Reliability	-
Validity	-
Reference	Schön, D., Anton, J. L., Roth, M., & Besson, M. (2002). An fMRI study of music sight-reading. *Neuroreport, 13*(17), 2285-2289.

Title	Naming of musical notes: A selective deficit in one musical clef
Author	Daniele Schön, Carlo Semenza, Gianfranco Denes
Year	2001
Purpose	• To investigate the ability of various tasks with a professional musician presenting a complex pattern of amusia after a left temporoparietal ischemia
Measurement	• Ability to transcode a notational code into a spoken code • Integrity of transcoding process leading from notational code to a motor performance • Ability to access clef information with same-different judgments on items in both clefs • Integrity of the note lexicon • Assessment of transcoding from verbal auditory code into a musical notation • Transcoding from verbal auditory code into motor performance • How the subject held musical notation in memory • Assessment of subject's notational ideograms
Assessments Used	*(See assessments used with neuropsychological examination)* Bock and Miller (1991) sentence completion task *(for nonmusical task)*
Description	
Subjects	1 subject: 65-year-old female professional organist with a cerebral ischemic lesion on the left parieto-temporal region (this means lack of blood flow to meet metabolic demand, resulting in cerebral hypoxia) For experimental tests, 3 professional musicians (average age of 49) with 30+ years of experience, who performed at ceiling
Instruments Used	Piano
Stimuli	**Oral Reading (solfeggio) and Instrumental Sight Reading at Keyboard** – 72 notes (36 in each clef), 60 sequences (30 in each clef). Total of 240 notes. Half were musical easily identified musical patterns (a single 'chunk'). All displayed on single staves with clef **Same-Different Discrimination** – "Strings" of 4 notes, one in bass, and other in treble clef, representing same or different sequences transposed at the octave. Presented on a sheet beside each other. 40 pairs, either: musically identical (same), graphically identical, structurally identical (i.e. contour and interval), totally different **Naming Played Notes** – 16 single notes and 56 sequences. The sequences could be musical or nonmusical, and were in high or low registers **Verbal Dictation** – 16 single notes and 30 sequences of 4 notes **Playing on Verbal Dictation** – 10 single notes and 16 sequences **Delayed Copying Task** – Same items from oral reading test **Naming and Writing Ideograms, Chords and Key Signatures** – In each session, 30 items involving clefs, rest, time signatures, dynamics, accidentals, slurs, and similar
Apparatus	*Finale 3.5* for presentation Paper and Writing Utensil
Method	*Neuropsychological examination: Assessed spontaneous language, used Italian version of Aachener Aphasie Test (Luzzatti, Wilmes and De Bleser, 1991), Reading, Digit span, Disyllabic word span, Spatial span (Corsi's Test), used Raven Progressive Coloured Matrices.* For more information, see Pg. 409-410. General music abilities: Music reading and related tests **Oral Reading (solfeggio)** – Two versions: single notes and groups of 4 notes. Subject read notes and sequences as presented **Instrumental Sight Reading at Keyboard** – Two versions: single notes and groups of 4 notes. Subject played notes and sequences as presented **"Same-Different" Discrimination** – Subject had to determine same-different from strings of 4 notes **Naming Played Notes** – Subject sat beside experimenter at piano, looked at

	experimenter's hand, then named the played notes Music Writing and Related Tests **Verbal Dictation** – Subject wrote dictation at a rate of approximately 1 note/s **Playing on Verbal Dictation** – Experimenter named notes and sequences and subject played them on a keyboard **Delayed Copying Task** – Single notes were presented for 1 s, and sequences were presented for 2 s. The subject could start notation when the presentation was **over** **Naming and Writing Ideograms, Chords, and Key Signatures** – In two different sessions, the subject was asked to name and write 30 musical items **Additional Nonmusical Tasks** – Sentence completion task, reading of complex numbers test
Results	Musical Abilities Tasks: Found rhythmic disturbance across tasks with melodic discrimination quite well preserved. Temporal Processing: Found that rhythmic discrimination was performed better than reproduction. Metric identification was unimpaired. Could not play anything from memory. **Oral Reading (Solfeggio)** – Subject performed better with G clef. No difference between musical and nonmusical items in both clefs. Overall impairment confined to oral reading in F clef. **Instrumental Sight Reading at the Keyboard** – Subject was 100% correct in single note task and 96.5% correct in sequence task **Same-Different Discrimination** – Subject correctly judged 36/40 pairs **Naming Played Notes** – At ceiling for single notes and 96% for sequences **Verbal Dictation** – With single notes, subject did not do well writing notes in F clef, with tendency to write them as if they were in G clef. With sequences, the F clef was also poor **Playing on Verbal Dictation** – Perfect performance **Delayed Copying Task** – Perfect performance with single notes. For sequences, G clef was 93.5% correct with musical items and 58.5% correct with nonmusical items. F clef was equally poor. This shows that subject was using a chunking strategy for G clef **Naming and Writing Ideograms, Chords and Key Signatures** – Subject able to name (20/22) and write (19/22) most of ideograms presented. Subject was unable to identify tonality on basis of accidentals, to write accidentals of a given tonality, to say whether a tonic triad was major, minor, or diminished. Therefore, only knowledge of musical ideograms was preserved **Additional Nonmusical Tasks** – Found that subject was able to successfully carry out linguistic production tasks heavily constrained by contextual, syntactic or syntactic-like information. Therefore, the failure to read the F clef orally is regarded as domain-specific
Reliability	-
Validity	-
Reference	Schon, D., Semenza, C., Denes, G. (2001). Naming of musical notes: A selective deficit in one musical clef. *Cortex, 37*, 407-421.

Title	Mind's ear in a musician: Where and when in the brain
Author	Martin Schurmann, Tommi Raij, Nobuya Fujiki, Riitta Hari
Year	2002
Purpose	• To study the temporospatial pattern of brain activity during auditory imagery
Measurement	• Brain activity specific to an auditory imagery task using magnetoencephalography (MEG) • Studying the time course of brain activation sites during auditory imagery in musicians
Assessments Used	-
Description	
Subjects	11 experienced professional-level musicians (ages 19-37 years) (Retained data for 10 of the subjects)
Instruments Used	-
Stimuli	Four conditions: Condition 1 - Visual control – Black dots on empty screen, Condition 2 - Auditory control – Piano sounds corresponding to notes g1, a1, h1, c2, Then these conditions presented in alternating blocks of 20 stimuli: Condition 3 - Note-sound associations - With 4 visual notes on staff with respective piano sounds, Condition 4 – Auditory imagery – With notes presented on staff but without auditory stimuli. For this last condition, subjects were instructed to imagine the corresponding sounds Additional control condition for 5 of the subjects: Attend-to-faces - Stimuli on staff but not notes. Instead, black-and-white drawings of faces. Aim of this to rule out that activation effects during auditory imagery would be due to attention to visual stimuli rather than auditory imagery task. Stimuli presented at irregular rate once every 1/5-2 s
Apparatus	MEG room with magnetic shielding Data projector for visual presentation Earpieces for auditory presentation 306-channel *Vectorview* whole-scalp neuromagnetometer (Neuromag Ltd., Finland) for recording MEG responses Surface electrodes above larynx (to make sure there was no vocalization)
Method	Subjects saw condition 1, then 2, then an alternation of conditions 3 and 4. After experiment, subjects were asked to give their own estimate of how well they could imagine the sounds (scale of 0-10). Additional control condition was presented between 2 and 3/4 conditions
Results	Condition 1 – Visual control – Activation in several posterior brain areas, Condition 2 – Auditory control – Main activation in supratemporal auditory cortices, Condition 4 – Auditory imagery – Activation in areas similar to those during visual control but also in further areas: first as enhanced activity in left and right occipital areas, then spreading to midline parietal cortex and extraoccipital areas (i.e. left temporal auditory association cortex and left and right premotor cortices) not activated during visual control condition. This implies a complex temporospatial activation sequence with multiple cortical areas, when musicians recall firmly established audiovisual associations
Reliability	-
Validity	-
Reference	Schurmann, M., Raij, T., Fujiki, N., & Hari, R. (2002). Mind's ear in a musician: Where and when in the brain. *NeuroImage, 16*, 434-440.

Title	The development of skill in reading music
Author	Lawrence Richard Scripp
Year	1995
Purpose	To investigate the questions: • Do changes in music reading skills provide evidence for new levels of cognitive development? • If so, what patterns of reading skill development emerge during first 2 years of undergraduate study • What are implications of a cognitive-developmental view of music reading skill acquisition for music educators and music psychologists? • To propose and test a cognitive-developmental model of music reading, with account of how skills progress from unstable, externally-dependent representations to stable, internally-generated representations of music notation
Measurement	To measure reading skill development, analyzing: measures of sightsinging, error detection skills, introspective comments • **Production** - Music reading errors is anything that deviates from the notation. Also considers error recovery. Possible errors: 1. Rhythm entrance error, 2. Rhythm figure error, 3. Rhythm recovery, 4. Pitch error, 5. Pitch recovery, 6. Error of omission, 7. Error of addition • **Error Detection** – 1. Undetected rhythm error, 2. Undetected pitch error, 3. 'Falsely' detected error • **Interview Questions** – 1. Self-assessment in terms of accuracy, 2. Self-awareness in terms of the learning processes, 3. Reports or descriptions of intrinsic qualities of and comparison between reading processes, 4. Internalization of musical imagery Scoring and transcribing team was graduate students, teaching assistants and alumni from New England Conservatory. Asked to apply professional standards while making subjective judgments
Assessments Used	Watkins-Farnum Performance Scale (WFPS) (Watkins & Farnum, 1954) – Stimuli adapted from this
Description	
Subjects	20 subjects selected at random from population of students required to take a 2-year course in sightsinging at New England Conservatory
Instruments Used	Voice, Piano (by experimenter)
Stimuli	**Sightreading Items for Sightplaying AND Sightsinging Task** - Adaptations from the WFPS and current sightsinging tests for first-year New England Conservatory students. Instrumental form of sightreading materials is more demanding with length, wider intervallic leaps, greater range of rhythmic variety. **Complete the Melody Task** – 3 responses: 1. Reading not informed by framework established in opening phrase, 2. Inferential reading of second phrase in accordance with context of first phrase, 3. Reading which accounts for tonal or rhythmic framework of first phrase, but which uses inferential understanding of that phrase to accurately read pitches or rhythms not expected in second phrase. **Perception: Error Detection Task** – 3 Conditions: 1. Identical with those in 'complete the melody' task, 2. Unexpected but musically appropriate notes added to second phrase, 3. Subject warned that items might contain 1 error, many errors, or no errors. **Interview Response** – Questions concerning accuracy and reading processes.
Apparatus	Video Recorder for analysis
Method	Proposes three dimensions of reading skill development - production, perception, reflection – as a cognitive lens into development Data from longitudinal sample collected 4 times over a 2 year period Entire session was a half hour in length **Sightplay/Sightsing** (Production) **Sightsing/Complete the Melody** – The first phrase is rehearsed until it is learned,

	and the second phrase is sightread. The examiner would show the first phrase, play it on piano, and allow for repeated rehearsal. Once learned, the subject would sing it one more time and proceed to the next section. One sample to orient subject to task. **Complete the Melody/Error Detection** – The experimenter plays the items once on piano and subject circles notes on page that do not correspond with the performance. The subject labels notes as either rhythm or pitch errors. **Interview Response** – Interspersed between production and perception tasks, questions facilitate reflective thinking while participant engaged in sequence of events of experimental session.
Results	Found that changes in music reading skills suggest new levels of cognitive development, but only when indicated by 1. Ability to sight-read music without an instrument, 2. Reflective understanding of music reading processes as flexible, problem-solving skills Found that sightsinging remains categorically different from sightplaying Found suggestion that reading skill development is a process characterized by gradual integration of cognitive skills, which indicates emergence of new forms of musical intelligence. Therefore, musical understanding at undergraduate level cannot be furthered by learning repertoire with voice or instrument alone Results of statistical and qualitative analyses confirm the developmental model of music reading: 1. Reading skill is comprised of 3 distinct and independent dimension of skill in production, perception, reflection, 2. Reading skill development can be assessed quantitatively and qualitatively for categorical changes within these dimension over time as well as for 'degrees of synchrony' across dimensions over time, 3. Musical development that emerges with reading skill acquisition characterized by a progression of 'stage-like' levels, 4. Evidence for musical development depends on analysis of qualitative differences in approach to reading tasks, in general characteristics of response to reading tasks, and in analysis of underlying cognitive structure of responses to reading tasks
Reliability	Production Tasks – 96-99%, Error Detection – 98%
Validity	-
Reference	Scripp, L. R. (1995). The development of skill in reading music. *Dissertation Abstracts International Section A: Humanities and Social Sciences*, 56 (6A), 2162.

Title	Distributed neural network underlying musical sight-reading and keyboard performance
Author	Justine Sergent, Eric Zuck, Sean Terriah, and Brennan MacDonald
Year	1992
Purpose	• To study the functional neuroanatomy of musical sight-reading and keyboard performance
Measurement	• Positron emission tomography (PET) for regional cerebral blood flow (rCBF) • Magnetic resonance imaging (MRI) for anatomical information of the subjects' brains
Assessments Used	-
Description	
Subjects	10 professional pianists
Instruments Used	Piano
Stimuli	Score of partita by J. S. Bach, not well-known
Apparatus	TV monitor for presentation Keyboard for performances *Philips Gyroscan* (1.5 T) for MRI imaging *Scanditronix PC-2048 Camera* for PET imaging
Method	Subjects participated in 7 task activation conditions. Main experimental condition: Each subject played the partita on a keyboard with the right hand while listening to the performance. Other conditions: 1. Visual fixation of the lit screen, 2. Listening to ascending and descending musical scales playing on a piano, 3. Playing ascending

	and descending scales on the keyboard with the RH while listening to what was played, 4. Presentation of a single dot in 1 of 5 quadrants of the screen and manual responses as a function of dot location, 5. Reading a musical score presented on a screen, 6. Reading a musical score presented on the screen and listening to its performance played on a piano
Results	Find the suggestion that sight-reading and piano performance entail processing demands that are realized by a cerebral network distributed over the four cortical lobes and the cerebellum. This parallels, but is distinct form, the neural substrates of verbal processing. The superior parietal lobe must participate in the spatial nature of musical notation reading
Reliability	-
Validity	-
Reference	Sergent, J., Zuck, E., Terriah, S., & MacDonald, B. (1992). Distributed neural network underlying musical sight-reading and keyboard performance. *Science, 257*(5066), 106-109.

Title	Effects of rote versus note presentations on rhythm learning and retention
Author	Patricia K. Shehan
Year	1987
Purpose	• To examine the effects of aural and visual approaches to rhythm reading and short-term retention • To determine effective avenues of rhythm reading in an attempt to understand the process of music literacy
Measurement	• Efficiency • Speed • Retention of learning rhythms
Assessments Used	-
Description	
Subjects	25 grade two and 24 grade six students with assumed minimal note-reading competence
Instruments Used	Woodblock
Stimuli	Author-composed four 8-beat rhythm patterns, consisting of quarter and eighth notes, one quarter rest, termination with a quarter note. All 4/4 time with patterns likely to be found in American folk songs. Maximum performance time of 4.5 s
Apparatus	Tape recorder for presentation of mnemonics and rhythm patterns Cards for visual presentation
Method	Quarter note = *Tan*, eighth notes = *Teka*, 2 quarter notes = *TON TON*. 4 modes of presentation of the rhythmic patterns: 1. Audio-rhythm (ar) (sounded on a woodblock), 2. Audio-mnemonics (am) (mnemonic syllables), 3. Audio-visual (vr) (illustration in bold block notation on a card, presented while pattern was sounded on a woodblock), 4. Audio-visual-mnemonic (vm) (combined notation and vocalization of the pattern). Each subject tested individually in a 15-minute session. The subject had a brief demonstration and explanation, and the first pattern was presented. After each rhythm presentation, the subject was to memorize and perform the rhythm pattern on a woodblock
Results	For beginning musicians, blending of visual and aural strategies may best facilitate the learning of rhythm patterns. Simultaneous use of auditory and visual channels appears to enhance learning retention of rhythms. Maturation may have an important role in learning the patterns faster, regardless of presentation mode
Reliability	-
Validity	-
Reference	Shehan, P. K. (1987). Effects of rote versus note presentations on rhythm learning and retention. *Journal of Research in Music Education, 35*(2), 117-126.

Title	The eye-hand span – An approach to the study of sight reading
Author	John Sloboda
Year	1974
Purpose	• To explore some of the similarities and differences between literacy reading and musical reading • To begin to look at how musicians read music and process it
Measurement	• Sight reading ability when slide was visible (counted number of errors) • Eye-hand span (EHS)
Assessments Used	-
Description	
Subjects	10 subjects; mostly music students. These subjects covered a wide range of sight reading ability, but of roughly equivalent age and musical experience
Instruments Used	Piano (did not say)
Stimuli	Experimental pieces were single line melodies written in medium range in the treble clef; little known English and French popular and folk melodies of the 1800s. There were no expression or phrase markings. Melodies were highly conventional and technically clear. Total of 15 melodies. Melodies were written out so that the size of notes and distance between notes was consistent.
Apparatus	Projector for stimuli Slides of each stimulus Metronome for tempo
Method	Subjects began to play each slide at its onset, preceded by a metronome beat. Each slide was displayed twice and turned off early in the piece and later in the piece. When the slide was removed, the subjects were to continue playing until they could not remember any more notes
Results	When the distance to the end of the phrase is 4 notes or less, all subjects were able to reach the phrase boundary. When the distance was 11 notes +, none of the subjects were able to reach the phrase boundary. The two differences in EHS are due to differing "coding strategies" employed by poorer and better readers. Found that the placing of the phrase boundary had a significant effect on the value of the EHS. It is clear that subjects retained considerable information from beyond the EHS, such as not playing the correct notes but playing the correct rhythm or contour of the melody.
Reliability	-
Validity	-
Reference	Sloboda, J. (1974). The eye-hand span – An approach to the study of sight reading. *Psychology of Music, 2*(4), 4-10.

Title	Visual perception of musical notation: Registering pitch symbols in memory
Author	John A. Sloboda
Year	1976a
Purpose	• To examine the way in which experienced musicians differ from non-musicians in their recognition of briefly exposed pitch notation
Experiment 1	
Measurement	• Musicians and non-musicians' ability to report individual notes presented very briefly, or musicians' ability for specific stimuli presented at longer durations
Assessments Used	-
Description	
Subjects	8 subjects; 4 were undergraduates of London University with little or no musical training; 4 were music students highly experienced in dealing with musical notation
Instruments Used	-
Stimuli	6 cards each had a musical stave of 5 lines extending across the entire card, with note heads (no rhythmic value). Cards had varying lengths of notes 1-6. 96 cards total: 16 for each array of notes. No array had the same note twice

Apparatus	*Cambridge* two-field tachistoscope for presentation of stimuli
Method	There were 10 practice trials. For all trials, subjects looked at each stimulus carefully and immediately following offset they attempted to record the sequence of notes (L to R) on response sheet. Cards presented in random order, in blocks of 24 with 1 min rest between blocks. There were two experimental conditions: 20 ms exposure duration and 2 s exposure duration. Before each stimulus, the subject was told the number of notes, and was told to produce the same number on the response sheet. Subjects went to the next card by pressing a thumb switch
Results	Strongly suggest the existence of dual coding mechanisms: A visual code operating at brief stimulus duration (e.g. 20ms). All subjects possess this code, though it is not efficient. Second code is non-visual and receives ample time to come into operation with a 2s exposure duration. Musicians can recall up to 6 notes. Non-musicians are disadvantaged with longer exposure and there is no efficient form of non-visual coding in which to retain notes
Experiment II	
Purpose	• To determine the precise interval after stimulus onset at which non-visual code begins to have an effect on musicians' performance
Measurement	• Interval of time from stimulus presentation to non-visual code • Number of notes (out of three) recorded correctly
Assessments Used	-
Description	
Subjects	8 subjects; 4 were experienced music readers (student age); 4 were musically naïve and students of London University
Instruments Used	-
Stimuli	36 cards (like those in Experiment I), but with three notes on each card
Apparatus	Three-field tachistoscope for presentation of stimulus. The first field held a "fixation stave", the second the card, and the third a visual mask of closely spaced dots and lines, to jumble the elements of the stimulus
Method	Before each trial, the experimenter gave a verbal signal, immediately followed by the fixation field for 2 s, and then the stimulus for the remainder of the trial. There were 6 durations: 20, 40, 60, 80, 100, 150 ms. The stimulus was followed by the mask for 2 s. The subject could write their response as soon as the mask appeared
Results	At 150 ms, musicians performed better than non-musicians, suggesting that it takes at least 100 ms for first item in stimulus array to be coded into abstract form by musicians. However, there is no evidence that musicians are using a naming code, and may be coding directly into a kind of underlying pitch representation
Experiment III	
Measurement	• To determine if interference of irrelevant tones affects musicians' coding of visual notes as tones • How musicians differ from non-musicians at this task • Number of notes correct out of 6 for each stimulus
Assessments Used	-
Description	
Subjects	10 subjects; 5 were musically naïve students of London University; 5 were experienced keyboard musicians of student age
Instruments Used	-
Stimuli	Cards as in Experiments I and II. 40 cards with 6 notes each, of a random sequence. Auditory interfering material included: 1. Male voice reading from textbook at normal conversational speed, 2. "Tonal" music (Bruch's violin concerto, Mvt I), 3. "Atonal" music (Boulez solo clarinet), 4. No auditory interference. Visual exposure conditions same as Experiment I
Apparatus	Speaker and tape deck for auditory stimulus
Method	The subject was told to reproduce the visual display while ignoring any background noise. Each stimulus exposure was 2 s. Interference conditions were in blocks of 5

	exposures, or 2 blocks for each condition. The subject had to produce 6 notes on each response card, as quickly as possible. Subject moved to next card by use of thumb switch
Results	Interference seemed to have no effect on performance of musicians, but either improved performance of non-musicians, except for speech effect. Musicians may have more devotion to music task at hand or they can better filter out extraneous sounds at a peripheral level
Experiment IV	
Measurement	• The subjects' retention of interference stimuli • To determine subjects' 80% performance level • To determine if two concurrent tasks might use the same code, and so interfere with each other
Assessments Used	-
Description	
Subjects	10 subjects; 5 were musically naïve students of London University; 5 were experienced keyboard musicians of student age
Instruments Used	-
Stimuli	The cards used for Experiment III. Interference materials: Paired sequences of letters; paired sequences of tones. For each, have the sequences were identical and half were different from each other by one item. For letters, one letter was easily confused with its partner in the other sequence, and for tones, one was altered by a semitone. Letters were from: B, C, D, G, J, K P, Q, T, V, Z. Tones were from the 11 consecutive pitches between D (above MC) to D flat one octave higher.
Apparatus	Electronic organ for performance Tachistoscope
Method	Subjects heard a sequence of letters or tones. After 5 s subject heard a second sequence, and had to judge whether it was the same or different from earlier sequence. During 5 s break, subject performed a 'filler' task, involving inking in 6 outlines of notes on a stave. The experimenter either read the letters live, or played the tones live. For each type of sequence, the tones and letters progressively got longer, in order to find the 80% performance level. There were 3 conditions: 1. Visual task alone (like in Experiment I), 2. Visual task with interfering letter memory task, 3. Visual task with interfering tone memory task. The interference task was an auditory sequence of tones or letters presented for subject to remember. Immediately following the last item the subject was to initiate the visual display with a thumb switch. The subject was to view the adaptation field of the tachistoscope while listening to the first sequence. After termination of visual stimulus (duration of 2 s), the subject had to record this stimulus as quickly as possible. Immediately following the last recorded note, the experimenter presented the auditory sequence for subject to make a same-different judgment. There were 10 practice trials, then 30 experimental trials, without break. 10 were visual task alone, 10 added letter-memory task, 10 added tone-memory task. Exposure duration was 2 s throughout
Results	No reciprocal interference found. Musicians either do not code visual note sequences by names/pitches, or they do code them by one of these systems but continue with concurrent activities with the same code simultaneously. Though the coding process is unknown in these experiments, they show that musicians have perceptual superiority when recording notes presented visually for durations 100 ms +
Reliability	-
Validity	-
Reference	Sloboda, J. A. (1976a). Visual perception of musical notation: Registering pitch symbols in memory. *Quarterly Journal of Experimental Psychology, 28*(1), 1-16.

Title	Phrase units as determinants of visual processing in music reading
Author	John A. Sloboda
Year	1977
Purpose	• To determine how structural and physical markers affect the eye-hand span of keyboard musicians.
Measurement	• Eye-hand span (EHS) with structural and physical markers • The interaction between structural and physical markers (These markers based on: Unitization hypothesis (supposes that groups can be delimited before any major analysis commences; segmenting by physical unit markers) and redundancy hypothesis (supposes that probably items require less perceptual analysis to be identified. Any kid of structure will create redundancy; structural elements will provide redundancy cues)) • The number of notes correctly recalled in unbroken sequence once the slide (with music) was switched off
Assessments Used	-
Description	
Subjects	6 subjects (ages 22-30); all accomplished keyboard sight-readers
Instruments Used	Electronic organ
Stimuli	36 8-phrase sequences of 9 matched sets of 4 sequences each. Phrase length was between 5-9 notes each. Matched set contained: a. Simple diatonic melody in folk or hymn melody style, with each phrase boundary marked by presence of a cadence and a relatively long interphrase space (physical and structural), b. Sequence identical to first in rhythm (same spacing), but with pitches assigned to values which made harmonic nonsense of sequence (physical), c. Sequence with comparable harmonic structure to a. but without interphrase spaces (structural), d. Sequence identical to c. in rhythm but with pitches assigned to value which made harmonic nonsense of sequence (no phrase markers)
Apparatus	Photographic slides of each sequence Projector for slides Screen for presentation Tape-recorder for analysis of presentations
Method	Subjects tested individually on organ, in front of projection screen. They were told that it was a test of sight reading ability with interest in the memory span of notes seen but not yet executed. The screen would be shut off and the subjects were to carry on playing until they had exhausted their memory of notes ahead. Subjects chose their own tempo, but had to begin playing with 5 s of the onset of the slide
Results	Physical and structural markers affect the EHS of experienced sight-readers. The two types of markers are independent, but structural markers are at least as important as physical markers. An increase in coincidence of EHS with phrase boundary is 27.8% for structural markers and 16% for physical markers. Structural marking always increases EHS, but physical marking may not increase EHS. To summarize, structural markers of phrases cause EHS to expand within an area of high redundancy (i.e. up to a phrase boundary). Physical markers of phrases can cause EHS to expand or contract according to whether the boundary is nearer or further than means span
Reliability	-
Validity	-
Reference	Sloboda, J. A. (1977). Phrase units as determinants of visual processing in music reading. *British Journal of Psychology, 68*, 117-124.

Title	Perception of contour in music reading
Author	John A. Sloboda
Year	1978b
Purpose	• To examine three possible attributes of differing detail: 'approximate position' of a note, 'absolute contour' (due to 2 notes), 'relative contour' (due to 3 notes). • To discover whether musicians retained more information than novices at any of these levels. • To see whether a superiority effect with exact position scoring might be obtained by varying the task requirements of the Sloboda (1976a) study
Measurement	• Relative contour (most global feature), then increasing in detail by absolute contour, approximate position, and finally exact position. Responses scored by exact position (1 if correct vertical position), approximate position (0 if in correct vertical position, and 1 if displaced by one vertical position on the staff, etc.), absolute contour (each pair of adjacent notes received score of 1 if vertical rank was preserved. Maximum score of each 4-note row was 3), and relative contour (each triplet of adjacent notes received score of 1 if direction of angle formed was preserved. Maximum score of each 4-note row was 2). Also, for each score, tested for 1. Whether musicians and novices differed, and 2. Whether there were significant differences between conditions
Assessments Used	-
Description	
Subjects	12 subjects from University of Keele. 6 were accomplished music readers, and 6 were musically illiterate
Instruments Used	-
Stimuli	144 stimulus cards, each with 2 music staves of 5 lines each. 48 of the cards were according to each of the 3 experimental conditions. Condition A and B: Stimulus cards were identical, differing only in order of presentation. Each had 8 notes (4 on each staff). Condition C: Cards had 4 notes only, but everything else the same as other conditions. On half the cars, the notes were on upper stave, and half on lower stave. For this condition, there were 12 examples of each of 4 types of configuration. Type 1: Straight-line contours. Type 2: One major direction change, like a rotated 'L'. Type 3: 2 major direction changes, like a rotated 'Z'. Type 4: No distinctive pattern with both spaces and lines used and note repetitions
Apparatus	*Cambridge* 2-field tachistoscope Response sheets and writing utensil
Method	Subjects held pen in writing hand to begin. Each condition began with 2 min rest and 5 practice trials. Subjects fixated on plain white field with small central fixation dot, which was ultimately a point midway between the 2 staves on a stimulus card. Then viewed stimulus for 50 ms. At end of exposure, stimulus was replaced by fixation field, and auditory signal was initiated for 500 ms. This signal was tone which took one of three pitch values (high, medium, or low) preselected for each trial. Condition A: Subjects were to report the upper row first when high tone sounded, and to report lower row first when low tone sounded. Tone always corresponded to critical row. Condition B: Subjects were to report upper row alone when high tone sounded, and to report lower row alone when low tone sounded. Tone always corresponded to critical row. Condition C: Subjects were to report single row presented as soon as they heard tone, which was always medium pitch. Subjects used response cards to reproduce as closely as possible the vertical position of notes, in correct left-to-right order. Asked to respond for each note even if it was a guess
Results	Absolute Contour Score: Resulted in significant difference for musicians vs. novices.

	Musicians retained more information about relative positioning of adjacent notes. Confirmed finding of Sloboda (1976a): Musicians are not superior to novices at reporting briefly exposed notation (with regard to exact-accuracy measure of performance). Musicians able to divide attention between staves, while novices did better when reporting top stave (means that novices use 'top down' strategy, similar to prose reading). Summary: 'Global precedes detailed' theory of perceptual professing accounts for perception of music notation. "In conclusion, it is a relief to discover that, when appropriate response measures are made, music reading does not constitute an exception to the rule that skilled performers are better than novices." (p. 331)
Reliability	-
Validity	-
Reference	Sloboda, J. A. (1978b). Perception of contour in music reading. *Perception, 7*, 323-331.

Title	Note-by-note music reading: A musician with letter-by-letter reading
Author	Maria Stanzione, Dario Grossi, Lucianco Roberto
Year	1990
Purpose	• To propose a music-reading model similar to a text-reading model, with two accesses – global and step-by-step, where reception and production of language is similar to that underlying reception and production of music, with a partial functional overlap. This is proposed while studying a musician with a left posterior temporoparietal lesion
Measurement	• Reading Tests: Reading ability of words and nonwords, reading time • Tests of Musical Ability: Reading time, especially as number of notes increase
Assessments Used	*From General Examination: Raven's progressive matrices (1947), Corsi's spatial span test (Grossi, Matarese, & Orsini, 1980), Wechsler digit forward test, "Barrage" test, Language examination of the "Centro di Neuropsicologia dell'Universita di Milano" (Basso, Capitani, & Vignolo, 1979), Token test (De Renzi & Vignolo, 1962)*
Description	
Subjects	A 26-year old patient; Italian-speaking music teacher. This subject suffered a head injury and later had right lateral homonymous hemianopia, spatial-temporal disorientation, vomiting, and headache. EEG revealed severe damage in left temporoparietal areas, with an intracerebral hematoma in the parietal and posterior parietal regions. The subject could not read text and musical scores well. 5 professional musicians also participated in 'Reading of Note Sequences' portion
Instruments Used	-
Stimuli	***Reading Tests:*** *Phonemic Spelling: low-frequency words of four letters, five letter, six letters, seven letters. 5 of each. Reading Test: 108 words, up to 8 letters each, and 108 nonwords, up to 8 letters each. Reading Time: 75 words and 75 nonwords of increasing length. 5 classes: 4-5 letters to 15-20 letters.* **Tests of Musical Ability:** General Music-Reading Test: Identification of single notes: 10 different notes on treble clef and 10 different notes on bass clef. Identification of time signatures and other musical symbols: musical scores with varying time signatures and other symbols such as sharps, flats, etc. Reading of musical scores: Musical scores with notes showing time values, based on time signatures of each score. Reading of Note Sequences: 5 groups of 10 measures, each containing 2-16 notes. Notes were from conventional piece of music, but unfamiliar scores
Apparatus	*Reading Tests: Cards for presentation of stimuli* (No other apparatus named)
Method	**Neuropsychologial Study** ***General Examination:*** *1. Raven's progressive matrices, 2. Memory Tests: Corsi's spatial span test, Wechsler*

	digit forward test, Verbal word span, 3. Spatial attention: "Barrage" test, 4. Agnosia: Color agnosia: Color naming, Color recognition, Picture agnosia: Figure recognition, Naming simple drawings, 5. Language examination: Spontaneous speech, Language examination of the "Centro di Neuropsicologia dell'Universita di Milano", Understanding with Token test, Writing, 6. Phonemic synthesis: Reconstruct whole word from single letters presented orally **Special Neuropsychological Examination** ***A. Reading tests****: Phonemic spelling, Reading test, Calculation of reading time* **B. Tests of Musical Ability:** General Music-Reading Test Identification of single notes: Subject named notes one at a time. Identification of time signatures and other musical symbols: Subject shown time signatures and other symbols and asked to identify. Reading of musical scores: Subject read notes of several musical scores, specifically to name the note values, based on the time signatures. Writing Music: Dictation: Subject took dictation from short pieces. Transposition: Subject wrote notes presented by examiner in other forms, such as transposing a note written in treble clef to bass clef, or transposing into different keys. Reading of Note Sequences: Subject read successions of notes aloud
Results	***General Examination:*** *Slow and difficult, though correct, reading. No linguistic, perceptual, or memory-related disturbances. Understanding and writing capabilities were intact, and good capability for phonemic synthesis. Therefore, the reading impairment was not associated with any other cognitive defect* ***Reading Tests:*** *Found normal reading capability. Only slight impairment with very long stimuli and particularly with nonwords. Subject could turn graphemes into phonemes. Reading time was much slower than control subjects. Reading time increased with length. Subject had letter-by-letter reading for both words and nonwords.* **Tests of Musical Ability:** General music tests showed no deficit. Found that time taken by subject to read music increased progressively with number of notes in a measure. This confirmed a disturbance in reading note clusters.
Reliability	-
Validity	-
Reference	Stanzione, M., Grossi, D., & Roberto, L. (1990). Note-by-note music reading: A musician with letter-by-letter reading. *Music Perception: An Interdisciplinary Journal, 7*(3), 273-283.

Title	Brain changes after learning to read and play music
Author	Lauren Stewart, Rik Henson, Knut Kampe, Vincent Walsh, Robert Turner, Uta Frith
Year	2003a
Purpose	• Focuses on music reading as a translational process between the encoding of a stimulus and the execution of a motor response • To test the hypothesis that music reading involves a sensorimotor translation in which the spatial information contained within musical notation is used to guide selection of the appropriate motor response. Looks at music reading in a learning context
Measurement	• Functional brain (fMRI) changes before and after the acquisition of music reading skill from scratch. This requires a unique experimental design and therefore uses two different tasks: explicit music reading and implicit music reading. Explicit needs different formats before and after training. Implicit uses an identical format before and after training • Skill acquisition: External music teacher examined each subject individually on keyboard skills, including scales, prepared piece from Grade 1 syllabus, sight reading. Subjects also completed a Grade 1 theory examination • Cognitive measure of music reading ability: Using musical Stroop task (Stewart),

	music reading ability skill measured indirectly by ascertaining degree to which musical notation for pitch interferes with required number to finger mapping • Behavioural: Reaction times pre- and post-training; Errors pre- and post-training for both music reading tasks. For explicit task, the cumulative response time for trials in which all keypresses were correct. For implicit task, simple reaction time and error percentages calculated for each trial type
Assessments Used	-
Description	
Subjects	2 groups of 12 subjects. Learner group had no music reading or playing experience, but agreed to 15 weeks of musical training (2 subjects had to have two extra lessons of 90 min). Nonlearner group was control group, had no musical experience, and was not offered training. First stage of selection was self-report musical training questionnaire. Second stage of selection was attempt to play a set of simple melodies on a keyboard using RH
Instruments Used	Piano
Stimuli	Keyboard tutorial book, music theory manual and exercise book for training Simple 5-note melodies with RH for task
Apparatus	*Yamaha PSS26* keyboard *2-T Siemens Vision* system (Siemens, Erlangen, Germany) used for echplanar (EPI) slices *Statistical Parametric Mapping Software SPM99* (Wellcome Department of Cognitive Neurology, London; http://fil.ion.ucl.ac.uk/spm) for data analysis
Method	Learners group required to attend a 90-min music lesson once a week for 15 weeks, in groups of 3-4. Lessons given by experienced music teacher and followed standard method with music reading taught in conjunction with practical keyboard skills. Also music theory taught, and all taught to Grade 1 level with the Associated Board, UK. Subjects in this group practiced regularly and completed music theory exercises During scanning, the following tasks were used: Explicit music reading task: Subjects produce series of keypresses in response to appearance of a sequence of 5 musical notes. Pre-training, with numbers; post-training, numbers replaced with nonsense symbols. Nonlearners used numbers during both tasks. Keypresses used thumb to little finger. Control and experimental blocks were 6 trials of 5 s each. Rest blocks (presentation of blank stave) were 3 trials of 5 each. Stimuli presented for 4.5 s. Subjects fixated on cross hair in center of screen for remainder of trial (0.5 s). Implicit music reading task: Based on implicit text reading paradigm (Price et al., 1996). Subjects indicated whether target (vertical line) was ascending or descending (above/below staff), using arbitrary up/down mapping to index and middle fingers. Control and experimental blocks were 20 trials of 1.5 s each. Rest block (presentation of blank stave) was 10 trials of 1.5 s each. Stimuli presented for 1 s. Subjects fixated on cross hair for remainder of trial (0.5 s). Subjects scanned with both tasks in 2 separate runs. (Subjects also performed 2 additional tasks, not reported in this study.) Control and experimental blocks were 30 s; rest blocks were 15 s. Each run was 9 min in total.
Results	Skill acquisition: All learner subjects reached standard of keyboard skill equivalent to Grade 1 (Associated Board, UK) over the 15 weeks Behavioural: Explicit: Subjects were faster to produce sequence of keypresses on control trials compared to experimental trials, both for reading by numbers (pre-) and reading by notes (post-). Implicit: No difference in reaction time or percentage error between experimental and control trials. No effect of training fMRI: Explicit: Training effect seen in superior parietal cortex. Implicit: Training effect seen in left supramarginal gyrus, left inferior frontal sulcus, and right frontal pole. Trial effect was restricted to post-training session. Nonlearner group showed no activation between experimental and control trials Summary: When learners used newly acquired skill of sight-reading, showed

	activation in bilateral superior parietal cortex. This is likely to reflect the specifically visuospatial translational element of music reading, and that such activation may be independent of skill level. Also found fusiform activity in explicit music reading task. Activation of left supramarginal gyrus in implicit reading task, in conjunction with Stroop interference seen after training, suggest that musical notation is automatically processed after 15 weeks
Reliability	-
Validity	-
Reference	Stewart, L., Henson, R., Kampe, K., Walsh, V., Turner, R., & Frith, U. (2003a). Brain changes after learning to read and play music. *Neuroimage, 20*, 71-83.
See Also	Stewart, L., Henson, R., Kampe, K., Walsh, V., Turner, R., & Frith, U. (2003b). Becoming a pianist: An fMRI study of musical literacy acquisition. *Ann. N.Y. Acad. Sci. 999*, 204-208.

Title	Becoming a pianist: An fMRI study of musical literacy acquisition
Author	Lauren Stewart, Rik Henson, Knut Kampe, Vincent Walsh, Robert Turner, Uta Frith
Year	2003b
Purpose	• To compare brain activation before and after training in musical notation and playing the keyboard
Measurement	• Functional magnetic resonance (fMRI) of brain areas
Assessments Used	-
Description	
Subjects	Musically naïve subjects (This is the only description)
Instruments Used	-
Stimuli	Explicit Music Reading Task: Series of sequences of 5 notes. Before training, there were numbers superimposed on the notes. After training, the numbers were removed Implicit Music Reading Task: A single vertical line extending above or below the 5 horizontal lines of the staff. For this there was a musical and nonmusical stimulus. The nonmusical stimulus was 'nonsense' symbols on a staff, while the musical stimulus used notation
Apparatus	*SPM99*, Welcome Department of Cognitive Neurology for analysis using statistical parametric mapping software
Method	90-minute music lesson once a week for 15 weeks, where subjects learned practical keyboard skills and music theory, to the Grade 1 level of the Associated Board, UK. During the scanning, the tasks involved were the explicit music reading task, and the implicit music reading task. Explicit Task had subjects producing series of keypresses in response to appearance of a sequence. Before training, subjects used superimposed numbers to lay sequences; after training, they relied solely on notation. Implicit Task was identical before and after training. Subjects indicated whether target was ascending or descending, using up/down mapping to index and middle fingers
Results	Found greater activation for experimental trials versus control trials, and for post-learning versus pre-learning. With "Learning to Play a Melody: Explicit Music Reading Task", training effect seen in right superior parietal cortex. The trial effect was significant before training but even greater after training. With "Effect of Exposure to Musical Notation: Implicit Music Reading Task", training effect seen in left supramarginal gyrus, left inferior frontal sulcus, and right frontal pole. Found that all voxels in these regions exhibited same relative pattern: trial effect was restricted to the post-training session. With "Activations Common to Both Explicit and Implicit Music Reading", inclusive masking revealed common training effects across explicit and implicit music reading tasks in bilateral superior parietal cortex, medial superior parietal cortex, left postcentral gyrus
Reliability	-
Validity	-
Reference	Stewart, L., Henson, R., Kampe, K., Walsh, V., Turner, R., & Frith, U. (2003b).

	Becoming a pianist: An fMRI study of musical literacy acquisition. *Ann. N.Y. Acad. Sci. 999*, 204-208.
See Also	Stewart, L., Henson, R., Kampe, K., Walsh, V., Turner, R., & Frith, U. (2003a). Brain changes after learning to read and play music. *Neuroimage, 20*, 71-83.

Title	Reading music modifies spatial mapping in pianists
Author	Lauren Stewart, Vincent Walsh, Uta Frith
Year	2004
Purpose	• To demonstrate that musical notation is automatically processed in trained pianists, in a vertical-to-horizontal visuomotor mapping
Experiment 1: Musical Stroop Task	
Measurement	• Effect of irrelevant musical notation on performance • Response times for executing the motor sequence under different conditions of number/note congruence • Errors – A trial was discarded if subject made 1 or more error • Response Time – The time taken to execute each entire sequence calculated • Cumulative Analysis – Response time data processed for each subject. Each stimulus was presented twice, so cumulative response time was averaged • Itemized Analysis – Data for first keypresses of every sequence considered separately from data corresponding to all second keypresses of every sequence, etc.
Assessments Used	-
Description	
Subjects	12 pianists from Royal Academy of Music; 14 nonmusicians from University College London
Instruments Used	- (Computer keyboard used to perform notes)
Stimuli	Musical “Stroop” Task: Each was a notated bar of 5 quarter notes (G, A, B, C, D), with numbers (1-5) superimposed on the note heads. Five types: 1. Baseline – had no musical notation; a row of five white numbers against a background black strip, 2. Congruent – musically congruent and spatially systematic (i.e. G = 1, A = 2, etc.), 3. Incongruent (Random) – musically incongruent and spatially unsystematic (e.g. B = 4, C = 1, etc.), 4. Incongruent (Systematic) – musically incongruent but spatially systematic (i.e. G = 5, A = 4, etc.), 5. Catch – part congruent and part incongruent (random). Total of 12 motor sequences. Each trial was presented twice
Apparatus	*Sibelius* notation software Laptop keyboard for performance *MATLAB* for stimulus presentation and recording
Method	Subjects sat with RH fingers resting over five adjacent keys of laptop keyboard. Practice session of 5 trials, then 120 test trials. Before each trial, central fixation point for 1 s. Then stimulus presented in center of screen for 3 s for pianist group, and 4 s for nonmusician group. Then fixation point reappeared for 1 s. Then the second stimulus appeared. Subjects told to ignore musical notation and use only the numbers to perform. Participants told to read from L to R, mapping the numbers to the respective fingers as quickly and accurately as possible
Results	Showed that the performances of musically literate pianists are affected by irrelevant musical notation, using number-to-finger mapping. Nonmusicians are unaffected by presence of musical notation. There is a relative interference effect for incongruent trials vs. congruent trials. Time taken to make first keypress was considerably longer than time taken to make subsequent keypresses. Musical notation is automatically processed in musically literate pianists
Reliability	-
Validity	-
Experiment 2: Nonmusical Stroop Task	
Measurement	• To investigate the nature of the representation of musical notation

	• Response time benefits and costs to making a response that was spatially compatible or incompatible with the learned stimulus-response mappings (used by pianists reading music)
Assessments Used	-
Description	
Subjects	8 pianists; 14 nonmusicians (all from experiment 1)
Instruments Used	-
Stimuli	Numbers presented at different spatial locations. Two versions used: 1. Vertical-to-horizontal stimulus-response task, 2. Horizontal-to-horizontal stimulus-response task. 5 numbers presented consecutively in 5 different locations. Types: 1. Baseline, 2. Congruent, 3. Incongruent (systematic). The 5 nonmusical stimuli were presented one by one, to prevent any strategic perceptual or oculomotor differences
Apparatus	Laptop keyboard for performance *MATLAB* for stimulus presentation and recording Computer monitor for presentation
Method	Subjects placed RH over 5 adjacent keys of laptop keyboard. They were told a number would appear in one of 5 locations and they were to map the number presented onto the appropriate finger and make the required keypress. They were to ignore any spatial information by the number's location. As soon as subject responded, the current stimulus disappeared and the next stimulus appeared. After 5 responses, there was a 1 s pause before the next trial
Results	The results support the hypothesis that musically literate pianists acquire vertical-to-horizontal stimulus-response mappings (through extensive practice) that generalize outside of the musical context. Both groups showed interference on horizontal-to-horizontal task, but had opposite patterns on vertical-to-horizontal task. Pianists possess set of stimulus-response mappings that correspond to the stimulus-response mappings required for reading and playing music
Reliability	-
Validity	-
Reference	Stewart, L., Walsh, V., & Frith, U. (2004). Reading music modifies spatial mapping in pianists. *Perception & Psychophysics, 66*(2), 183-195.

T

Title	Sources of individual differences in music sight-reading skill
Author	William B. Thompson
Year	1985
Purpose	• To investigate the contribution of 2 factors to sight-reading skill: 1. Music encoding efficiency, 2. Degree to which encoding and performance operate in parallel
Measurement	• Eye performance span, using light-out technique (Levin, 1979; Sloboda, 1974) • Sight-reading, with average inter-note interval obtained by dividing time of sight-reading piece by 96 (96 notes in piece) • Choice reaction time • Music Encoding (through recall task), scoring number of correctly recalled notes. Used absolute criterion where only notes that were correct in pitch and order were scored • Letter Encoding, scoring identically as above • Choice Reaction Time, scored reaction of performance to the nearest 10 ms
Assessments Used	Watkins-Farnum Performance Scale (WFPS), Form A (Watkins & Farnum, 1954)
Description	
Subjects	30 flute players; ages 17-31. Recruited from psychology classes and local community
Instruments Used	Flute
Stimuli	WFPS: Form A Sight-reading piece: Single piece of randomly generated music: C, D, E, F, G, A (one octave above MC) randomly arranged in 24 measure piece, 4/4 time, with each note occurring 16 times Eye-performance Span: 7 selections from Gekeler (1969); moderate difficulty; in major keys with no more than 3 flats or sharps; average of 30 measures (20-41 measures) Music Encoding Test: Slide of 2, 4, 6, 8 notes in treble clef. Stimuli chosen from Gekeler (1969) and Voxman (1954). 5 of each type of stimulus shown in random order; total of 20 trials Letter Encoding Test: 4, 6, 8, 10 letters; consonants only in random order, with no repeated letters
Apparatus	Tape recorder for performances Slides for presentation of stimuli Video monitor for presentation of stimuli for choice reaction task
Method	Subjects performed 6 tasks: WFPS, sight-reading random music, eye-performance span, music-encoding, letter-encoding, choice reaction time. WFPS: Form A was divided into 2 sections, played at different times in the experiment Sight-Reading: Subjects sight-read a single piece Eye-performance Span: Used light-out technique, where music is suddenly and unexpectedly removed during performance. Music removed at 3 different places for each of the 7 pieces. Subjects were to play only the notes (after offset of slide) that they knew were present. Light-outs did not occur during first 3 measures or within 3 measures after a previous light-out Music Encoding Test: The 20 trials were shown in random order, with 5 strings of each length shown for 2 s. Immediately after each trial, subject wrote notes on music paper as accurately as possible, with unlimited time for recall Letter Encoding Test: Identical to above test Choice Reaction Time Task: Subjects played the indicated pitch (1 of 6 possible pitches; C-A, above MC) as soon as they saw it displayed on video monitor. 10 practice trials followed by 30 randomly order trials (5 for each pitch) Music-Background Questionnaire administered, followed by each of the 6 tasks in

	reverse order with a different set of materials Total number of trials: WPFS, 2; Sight-Reading, 2; EPS, 42; Music Encoding, 40; Letter Encoding, 40; Choice Reaction Time, 60
Results	Found that sight-reading ability was highly correlated with EPS and music-encoding. Sight-reading skill also correlated with choice reaction time to individual notes. This suggests that EPS increases when later processing steps are more efficiently coordinated (not just encoding efficiency). The parallel processing concept may better represent the sight-reading process. This is made up of 3 steps: 1. Music notation transformed into internal representation (encoded), 2. Internal representation is translated into motor commands for execution, 3. Commands are executed and transcription process is complete. These steps in sight reading can be performed at the same time, and the degree to which they operate in parallel is "major source" of individual difference in skill
Reliability	-
Validity	-
Reference	Thompson, W. B. (1985). Sources of Individual Differences in Music Sight-Reading Skill. (Doctoral dissertation, University of Missouri – Columbia, 1985). *Dissertation Abstracts International,* B 47 (02), 828.

Title	The study and evaluation of certain problems in eartraining related to achievement in sightsinging and music dictation
Author	Marvin S. Thostenson
Year	1967
Purpose	• To estimate the growth in dictation proficiency achieved in the course of the first year of ear training • To estimate the relationship between proficiency in music dictation and in sightsinging at graduate and undergraduate levels • To note the interrelationships between these proficiencies under varying conditions associated with the music degree, major instrument, and amount of training in years of private study
Pilot	
Purpose	• To validate tests of sightsinging and music dictation
Measurement	• Competence in pitch, rhythm, and in combination Four people administrated and evaluated the test
Assessments Used	Author-composed CSS76 Criterion Sightsinging Test Author-composed PRM78, Pitch, Rhythm, Melodic Dictation Test
Description	
Subjects	54 University of Iowa students
Instruments Used	Voice
Stimuli	Initial sightsinging test: 1. 12 simple intervals, 2. 24 4-note pitch phrases without rhythm, 3. 20 2-bar rhythm phrases without pitch, 4. 20 2-bar phrases with pitch and rhythm The CSS76 Criterion Sightsinging Test: 76 items, with only treble clef and a pitch-range of 2 octaves The PRM78, Pitch, Rhythm, Melodic Dictation Test: 78 4-foil multiple-choice items, including, 1. 30 short phrases with unchanging rhythm, 2. 24 short phrases with unchanging pitch, 3. 24 short phrases with differing pitch and rhythm content
Apparatus	-
Method	The investigator began with a 30-minute sightsinging test. PRM78, Pitch, Rhythm, Melodic Dictation Test: Subjects were to listen to the tape recording of the test and select the foil whose notation correctly represented the sounds played
Results	The first administration of the test yielded positive reliabilities
Reliability	.84 and .94 (inter-correlation coefficient of .86)

Validity	-
Main Experiment	
Purpose	• To estimate the correlation of the separate sections of the PRM78 Dictation Test with proficiency in the CSS78 Criterion Sightsinging Test • To estimate the correlation of the PRM78 with proficiency in the CSS76 • To assess the growth in music dictation achievement resulting from formal study of ear training in the first year • To compare the average achievement levels in music dictation and sightsinging attained by first-year and non-first-year students Secondary purpose: To study other relationships developed logically from collected data under three conditions: Interrelationships between proficiency in music dictation and sightsinging and 1. Type of degree, 2. Category of major instrument, 3. Amount of private instruction on major instrument
Measurement	• Data analyzed for central tendency, dispersion, inter-correlations between sections and total tests, and for reliability
Assessments Used	Author-composed CSS76 Criterion Sightsinging Test Author-composed PRM78, Pitch, Rhythm, Melodic Dictation Test
Description	
Subjects	First-year university students of 5 college and university schools of music (for PRM78) First-year university students of 4 college and university schools of music (for CSS76). Administered individually 97 graduate students; 36 special graduate students in summer workshop; 15 sophomores in fourth semester of ear training (for PRM78 and CSS76 tests) 4 other groups of students (no details given) provided more data (for PRM78)
Instruments Used	Voice
Stimuli	Author-composed CSS76 Criterion Sightsinging Test Author-composed PRM78, Pitch, Rhythm, Melodic Dictation Test. Later, 78 items reduced to 72 items
Apparatus	-
Method	PRM78 Dictation Test given early in first semester, and at end of year students were retested CSS78 administered to students by graduate assistants near end of school year
Results	Graduate students excelled in pitch concepts Undergraduates were challenged by rhythm Little difference between those seeking the two different degrees, but graduate students were superior, especially with sightsinging A small group of string majors had higher scores than other instrumental groups on both tests. Keyboard majors also had somewhat higher scores. Vocal majors had lowest scores The rhythm factor of the tests is more difficult to test than pitch
Reliability	PRM78 = .88 (pretest and retest) for undergraduates, = .90 for graduate students. Also, .86 (pitch), . 68 (rhythm), .73 (melody) CSS76 = .95. Also, .84 (singing intervals), .93 (singing pitch phrases), .82 (singing rhythm phrases), .92 (singing melodic phrases)
Validity	-
Reference	Thostenson, M. S. (1967). The study and evaluation of certain problems in eartraining related to achievement in sightsinging and music dictation. *Bulletin of the Council for Research in Music Education, 11*, 14-35.

Title	The perceptual span and the eye-hand span in sight reading music
Author	Frances E. Truitt, Charles Clifton, Jr., Alexander Pollatsek, Keith Rayner
Year	1997
Purpose	• To examine the perceptual span of pianists by determining what the size of the perceptual span is for a task, or from what region of a visual display is useful information acquired on each fixation
Measurement	• Perceptual span, by use of a moving-window technique, called the eye-contingent moving window paradigm • Eye-hand span • Each keypress by the subject, and its duration • Total playing times • Eye-movements: Average fixation duration, number of progressive and regressive fixations, average progressive and regressive saccade length.
Assessments Used	-
Description	
Subjects	8 pianists of "average to above average ability". Average piano experience was 10 years; average of formal piano lessons was 7 years
Instruments Used	Piano
Stimuli	Simple melodies, played with one hand. Short piano pieces taken from beginning of *Mikrokosmos, Vol. 1* (Bartók, 1940). 5-note range with no accidentals. 32 melodies were used. All in 4/4 metre and with only quarter notes and quarter rests, and half notes and half rests. 13 of the melodies also had ties and/or phrasing. Ranged from 9-18 measures in length, or from 2-4 lines. Only one of the staves from the original music was used. Each melody divided into non-overlapping 2-beat regions; 2 regions per measure. No-window condition: Entire melody displayed on screen throughout the trial. 2-beat, 4-beat, 6-beat conditions: Fragments of the melody were added to the display as eyes moved through the region. Blank staves were displayed to right of the window, with melody displayed in all regions previously fixated. Display changes were completed in maximum of 15 ms after a new fixation was detected
Apparatus	*Noteworthy Composer* for stimuli *Fourward Technologies Dual Purkinje Image Eyetracker* with bitebar *Yamaha* keyboard Tactile markers to show the C position on the keyboard 17" VGA monitor, 80 cm from eyes, for presentation Filter on screen to decrease brightness MIDI interface between computer and keyboard for recording (CMT: Carnegie-Mellon Music Tools) Metronome for some of performances
Method	Subjects arrived and bitebar was prepared for stabilizing their head. They were seated and adjusted for comfort. The eye-tracker was calibrated. Examiner told subjects which hand would be used for each melody, and with which finger they would begin. Part 1: Subjects played through 16 of 32 melodies with metronome. This was for training and these results not used. Part 2: The moving-window technique was used with the 3 restricted window conditions, and a condition with no window. Subjects played all 32 melodies without metronome. They were instructed to maintain performance as close as possible to metronome tempo from training.
Results	Found 1. Average eye-hand span (EHS) is small (>1 beat for less-skilled group and 2 beats for skilled group), 2. Perceptual span is also quite small (most measures affected only if window was restricted within 2 beats of fixation). Therefore, subjects usually fixating close to note they were playing. Average EHS was slightly over 1 beat. Average perceptual span appeared to be >2 beats but <4 beats. Skilled subjects may

	have extracted information 5 beats ahead of the hand
Reliability	-
Validity	-
Reference	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.

Title	Eye movements in a simple music reading task: A study of expert and novice musicians
Author	Andrew J. Waters and Geoffrey Underwood
Year	1998
Purpose	• To determine the effect of tonal complexity of stimuli on task performance and eye movement behavior • To determine if there was a difference in task performance and eye movement for expert and novice musicians
Measurement	• Eye movements of expert and novice musicians • Measure of discriminability and bias were computed for each subject in each condition Signal Detection Theory was used to analyze discrimination performance
Assessments Used	-
Description	
Subjects	22 musicians; 11 were 'expert', with 10+ years of formal training, and all had high achievement with at least one instrument with a treble clef register. 11 were "novice", with 2 years or less of training, but had some knowledge of notation
Instruments Used	-
Stimuli	20 "tonally simple, visually simple" excerpts written. All had a treble clef and then 4 notes, and were either a simple scale or arpeggio. All notes were in a major diatonic scale. Each excerpt had two or less accidentals, and had one without changes in contour 20 "tonally simple, visually complex" excerpts arranged from original stimuli. Each of these had 2 contour changes 20 "tonally complex, visually simple" excerpts created by a shift of one or two notes from the original stimuli. This made them no longer part of a diatonic scale 20 "tonally complex, visually complex" excerpts created by using tonally complex structures and arranging them with 2 contour changes All of the above stimuli were divided into two sets of 40; one set for "same" trials and one set for "different" trials. Also, 16 practice trials were created
Apparatus	Eye Tracker monitoring equipment, where the horizontal position of the eye is detected by the reflection of an infrared beam from the cornea onto a matrix. The eye position is sampled every 4 ms. Eye Tracker recorder is mounted to subject's head High-resolution monitor for stimulus presentation Push-button response box for subjects to record answers Microcomputer for timing
Method	Subjects were told that half the trials would be the "same" and half would be "different". "Different" meant that each note position would be equally likely to differ. They were instructed to respond quickly and accurately. The Eye Tracker was put on the subject and calibrated. Subjects did practice trials, then 2 blocks of 40 trials, with a break for re-calibration between. The 2 stimuli were presented for 800 ms, one after another, with a fixation cross before each
Results	Experts performed more accurately and rapidly than novices, with clear quantitative and qualitative differences Exerts responded more accurately and rapidly to tonally simple stimuli; Novices showed no difference Experts had more fixations on the first stimulus presentation Experts used less fixations overall before responding
Reliability	-
Validity	-
Reference	Waters, A. J., & Underwood, G. (1998). Eye movements in a simple music reading task: A study of expert and novice musicians. *Psychology of Music, 26*(1), 46-60.

Title	Studying expertise in music reading: Use of a pattern-matching paradigm
Author	Andrew J. Waters, Geoffrey Underwood, John M. Findlay
Year	1997
Purpose	• To investigate perceptual processing of music notation • To introduce a pattern-matching paradigm that can directly test hypothesis that skilled sight-reading is associated with ability to rapidly process groups of notes • To suggest how paradigm can be developed for future study into perceptual processing skills in music reading
Measurement	• Response time • Eye movement
Experiment 1: Pilot	
Purpose	• To determine the skill sensitivity of the pattern-matching paradigm • To investigate the effect of the structure of the stimuli of task performance
Measurement	• Processing of pitch and duration simultaneously
Assessments Used	-
Description	
Subjects	3 groups: 2 'expert' groups (achieved Associated Board Grade 8 in monophonic instrument) and a novice group (no advancement beyond Grade 2)
Instruments Used	-
Stimuli	Use of a "same-different" matching task. 2 visually presented musical strings. All stimuli composed of 10 notes (2 measures with 5 notes each), all in 3/4 or 4/4 time. No key signature, no accidentals. There was a total of 5-8 different pitches and 3-5 different duration values Of the originals, each was randomly assigned to 'same', 'pitch different', 'duration different' condition For pilot, there were 4 types of stimuli: temporally coherent, pitch coherent; temporally coherent, pitch randomized; temporally randomized, pitch coherent; and temporally randomized, pitch randomized 30 melodies (known as 'originals') composed
Apparatus	Apple Macintosh computer and keyboard Psychlab for stimulus presentation, control, and recording
Method	Each subject tested in a cubicle They were told that there would be same and different trials The subject compares the 2 stimuli quickly and accurately, and presses one of two response buttons (either same, or different). They were to press one button with RH for same response, and another button with LH for different response For same trials, the pitch and temporal structure is identical. For different trials, there is a small alteration in either pitch of one notes or duration of one note. Alteration could be anywhere in stimulus They saw a fixation cross for 2 s to left of center on screen, then stimulus presentation, with clef were cross has been. Stimulus remained until response was made. Inter-stimulus interval of 3 s. Started with practice trials, then 2 experimental blocks of 60 trials each, with 10 min break between
Results	Found clear quantitative and qualitative differences in task performance: expertise effects for RTs, but no large difference in accuracy. Also expertise effects for RTs of 'different' trials. Novice groups showed no sensitivity to temporal structure for temporally randomized stimuli The pattern-matching paradigm was validated through this pilot study
Reliability	-
Validity	-
Experiment 2: Main Study	
Purpose	Eye movement recordings to provide explanation for expertise effect in pilot study • To investigate effect of temporal randomization of stimuli on comparison

	behavior of experts • To determine effect on performance of pattern separation
Measurement	• Eye movements: Horizontal signal from one eye and vertical signal for other eye, sampled every 5 ms • Response times • Response choice
Assessments Used	-
Description	
Subjects	3 groups: Group 1 – 8 full-time music students from University of Durham. Group 2 – 8 psychology students with music experience on monophonic instrument. Group 3 – 8 nonmusicians (age range for all = 18-29 years)
Instruments Used	-
Stimuli	60 melodies generated and randomly divided into sets A and B. For each melody, temporally randomized counterpart generated. The 60 stimuli randomly split into 3 groups of 20 each, corresponding to 'same', pitch-'different', and duration 'different' trials. 2 trials for each stimulus – one for 'near' and one for 'far'. 'Near' and 'far' patterns: 'Near' separated by 2.5 cm (as in Pilot Experiment), and 'far' separated by 7.5 cm Set A – near, Set A – far, Set B – near, Set B – far 2 practice blocks of 12 trials each (for near and far) Subjects did not see same stimulus twice (in near and far conditions)
Apparatus	Eye movement recorder: Differential limbus reflection technique and a binocular infrared system with the Skalar IRIS system (description by Reulen et al., 1988), with chinrest Deluxe Music for stimuli generator Apple Macintosh Quadra for presentation Display and Record presentation package for timing of responses, control, and recording
Method	Each subject received instruction sheet, and then calibrated (2 calibrations) eye movement apparatus, inserted on head Procedure same as in Pilot Experiment Session lasted 40 minutes
Results	*For Response Data (reaction times and accuracy measures):* No large differences in overall accuracy, but large differences in RTs. Groups 1 and 2 much more rapid in performing comparisons More errors made on duration-'different' trials than pitch-'different' trials for Groups 1 and 2 Groups 1 and 2 showed more sensitivity to temporal structures than nonmusicians More errors made with near trials, though compared faster, therefore speed/accuracy trade-off *For Eye-Movement measures:* Groups 1 and 2 used fewer flips (MC: What does 'flip' mean?). Assumed that these groups used larger units to compare stimuli in 'same' trials. These groups used less viewing time to process comparison units, and fewer fixations to process units. Found briefer fixations with greater skill level For near/far trials, subjects used longer unit views for far trials, but with more fixations to encode the units Experts use longer units
Reliability	-
Validity	-
Reference	Waters, A. J., Underwood, G., & Findlay, J. M. (1997). Studying expertise in music reading: Use of a pattern-matching paradigm. *Perception & Psychophysics, 59*(4), 477-488.

Title	A multimodal neural network recruited by expertise with musical notation
Author	Yetta Kwailing Wong and Isabel Gauthier
Year	2010
Purpose	• To identify the loci of visual perceptual expertise for musical notation • To test whether the right transverse occipital sulcus (TOS) is recruited with expert music-reading skills by attempting to replicate the exclusive activation of the right TOS for music-reading experts as done in an earlier study (Nakada et al., 1998) • To see if single notes or short sequences could automatically engage a distributed multimodal network • To compare specialization for single notes and 5-note sequences to test the counterintuitive prediction that single notes may elicit stronger and more widespread expertise effects in the visual system to test whether the neural response to notes in these areas (that is, a network of visual and non-visual areas) would correlate with the degree of visual skill when perceiving musical notation
Measurement	• Pretest: Subjects wrote down the letter name of as many musical notes as possible within 2 minutes from a randomly generated music score • Measuring expertise effects during visual judgments with musical notation, in both visual and non-visual areas of the brain • fMRI • Perceptual fluency • MRI imaging
Assessments Used	-
Description	
Subjects	10 music-reading experts (8 women, mean age = 21.4) and 10 novices (7 women, mean age = 23.8). Experts had at least 10 years of reading experience. Novices had very limited reading experience. The subjects were divided into two cohorts, each with 5 experts and 5 novices
Instruments Used	-
Stimuli	Cohort 1:36 pictures in each of 8 categories of objects – faces and common objects in gray-scale, single letters, single mathematical symbols, and single musical notes in black-and-white. The letters were in three fonts. The mathematical symbols were in Times New Roman font. The musical notes included 9 different notes on the 5-line staff, from D below the bottom line to G above the top line, in 4 different time values (half notes, quarter notes, eighth notes, sixteenth notes). Also used were 5-letter consonant strings, 5-mathematical symbol strings, 5-note sequences. The letters were in Courier font. The notes always included 3 quarter notes and 3 eighth notes. All strings were randomly generated. Cohort 2: Identical stimuli except for the following – stimuli were of different fonts and slightly different sizes. All single stimuli and all string stimuli were presented on an identical 5-line staff background. In the different version of all stimuli there was a small gap in one of the 5 staff lines. The gap was randomly located anywhere on the staff. For strings, the gap was always on the top or bottom line
Apparatus	*Apple* computers for experiment *Matlab* and *Psychophysics Toolbox* extension LCD panel for projection of stimuli Mirror for subjects to view the stimuli *3-T Philips Intera Achieva* scanner for imaging
Method	fMRI Task: Two types of experimental scans. The first was the single run, including blocks of single notes, single letters, and single mathematical symbols. The second was the string run, including blocks of note sequences, letter strings, and mathematical symbol strings. There were 3 runs of each, each lasting 5 minutes and 36 seconds and consisting of 18 16-second blocks (6 blocks for each stimulus condition) with 3 16-second fixation blocks interleaved at regular interval for

	baseline. Also used were 2 localizer runs showing blocks of faces, objects, single letters, and single mathematical symbols for localizing the fusiform face area and letter-selective area(s). Each localizer run was 5 minutes and 20 seconds and consisted of 16 16-second blocks of testing stimulus (4 blocks for each stimulus condition) and 3 16-second fixation blocks appeared regularly. In all runs, a 10-second fixation block and a 6-second fixation block were added at beginning and end. Cohort 1: Performed a one-back task where they pressed a key with their right index finger as fast as possible if they detected immediate repeats of a stimulus. Cohort 2: Performed a gap detection task where they pressed a key with their right index finger as fast as possible when they detected a gap on any of the 5 lines. For both cohorts during localizer run: Pressed a key if they detected immediate repeats of stimulus. Measure of Perceptual Fluency: (Only 19 of the subjects performed in this task) There was a fixation cross in center of screen for 200 ms, followed by 500 ms premask, then a target 4-note sequence for a varied duration. After a 500 ms postmask, 2 4-note sequences appeared side by side, one identical to target sequence and the other shifted by one step. Subjects had to select the identical sequence and respond by keypress. Also identical for letter strings
Results	Found that a widespread multimodal network of brain regions has higher selectivity for musical notation in music-reading experts than novices. The network includes visual, auditory, audiovisual, somatosensory, motor, parietal, and frontal areas. Visual stimuli as simple as single notes or 5-note sequences are enough to activate this distributed network. It is suggested that the recruitment of the areas is relatively automatic. With experts, there are expertise effects in the bilateral early visual areas, bilateral fusiform gyrus, and inferior temporal areas. They failed to replicate the expertise effect for the reading of musical scores as reported by Nakada et al., 1998. Found that the TOS responded similarly to notes and mathematical symbols in both groups. This suggests that the area is not specially recruited for musical notes per se or for expert perception of musical notation. No observation of hemispheric asymmetry in processing of notes. Expert perception of notation recruited ventral temporal areas in both hemispheres. With expertise in musical notation, the authors note that they lack behavioral studies that would characterize the factors that distinguish their processing from other kinds of object recognition. They speculate that visual-motor transcoding could be more important for musical notation than other objects, as notation is use as a cue for motor execution
Reliability	-
Validity	-
Reference	Wong, Y. K., & Gauthier, I. (2010). A multimodal neural network recruited by expertise with musical notation. *Journal of Cognitive Neuroscience, 22*(4), 695-713.

Title	Investigations into the role of early visual cortex in expertise reading musical notation
Author	Yetta Kwailing Wong
Year	2010
Purpose	• ERP Experiment: To examine the temporal dynamics of the visual selectivity for musical notes with music reading expertise
Measurement	• Visual expertise effects • Electroencephalogram (EEG) • Electrooculogram (EOG) for horizontal eye movements • Perceptual fluency measured to quantify music-reading expertise
Assessments Used	-
Description	
Subjects	Expert Group: 11 subjects (including the author) with at least 10 years of music reading experience (7 females, mean age = 21.7) Novice Group: 11 subjects reporting being unable to read music (6 females, mean age = 25.0)
Instruments Used	-
Stimuli	Single musical notes, single Roman letters, or single pseudo-letters. All stimuli were either on a 5-line staff or not. 18 black-and-white images in each of the 3 objects categories. Notes ranged from E on the bottom line to F on the top line in 2 different time values (quarter notes and sixteenth notes). Roman letters were 18 uppercase letters in Courier font. 18 pseudo-letters created by various combinations of parts form Roman letters. For no-staff stimuli, 6 musical notes were used (quarter note, eighth note, sixteenth note, either pointed upward or downward)
Apparatus	*Mac Mini* for experiment *Matlab* with *Psychophysics Toolbox* extension Monitor for viewing Gamepad for pressing key Sensor cap for EEG recording
Method	The stimuli were presented briefly one after another. Subjects were required to detect immediate repetitions of a stimulus (a one-back task). Each of the 6 object categories was presented in blocks of 6 trials. Each block began with fixation dot at center of screen for 500 ms, followed by the 6 trials, each with a stimulus presented for 700 ms and then fixation dot for 250-450 ms. Subjects pressed a key as fast as possible when they detected a repeat of the stimulus
Results	Perceptual fluency: Experts had higher perceptual fluency than novices for music sequences but not for letter strings Behavioral results of ERP study: Accuracy was better for Roman letters than for notes. No other significance ERP results: For Musical Notation (on-staff): Expertise effect... ...for C1 – more positive for experts. Suggests that the C1 is selective for notes with expertise ...for N170 – expertise effect for notes were obtained in both hemispheres for notes on the staff ...for P3 – no expertise was found for noes on the staff ...for CNV – expertise effect for musical notes suggest that anticipation for notes is altered by music reading expertise For Musical Notation (no-staff): Expertise effect... ...for C1 – expertise effect was observed for notes without staff, suggesting that higher sensitivity for notes is not limited to individuation of notes or pitch processing of the notes, and may be related to perceptual fluency with shape of the notes ...for P3 – no expertise effect was found ...for CNV – none found For Letters (on-staff): Selectivity observed for N170 and P3

	For Letters (no-staff): Selectivity only observed for P3
Reliability	-
Validity	-
Experiment 2	
Purpose	• To investigate whether music reading expertise alleviates crowding in the parafoveal visual field • To relate the crowding effect to the ERP expertise effects from experiment 1
Measurement	• Measured basic visual functions with two kinds of crowding: 1. The target note and its line could be flanked vertically by 4 extra lines, 2. The target note and its line could be flanked horizontally by 2 extra notes • Crowding was measured with a set of control stimuli • Measured far and near acuity and contrast sensitivity • Perceptual fluency and holistic processing of musical notes
Assessments Used	-
Description	
Subjects	All subjects from experiment 1 except the author
Instruments Used	-
Stimuli	Black in color presented on grey background, presented for 100 ms randomly on left or right. All stimuli had black elliptical dot either on, above, or below the middle horizontal line
Apparatus	*Mac Mini* for experiment *Matlab* with *Psychophysics Toolbox* extension Monitor for viewing Gamepad for pressing key Sensor cap for EEG recording
Method	Subjects judged whether a black dot was presented on a line or on the space above or below the line. A central fixation dot was shown for 500 ms, followed by stimulus for 100 ms. Subjects judged with keypress. They were first tested with musical stimuli followed by the control stimuli
Results	Perceptual fluency: Experts had higher perceptual fluency than novices for music sequences but not for letter strings Basic visual functions: All participants had normal far and near acuity, and all but one novice had normal functional contrast sensitivity Crowding: With extra lines, experts performed better than novices for both baseline and crowded conditions. With flanker notes, experts and novices performed similarly for baseline condition, but crowding effect was smaller for experts than novices. For control stimuli, no main effect or interaction found. Summary: experts experienced less crowding than novices when crowding elements were staff lines or flanking notes. Suggests that music reading experience helps alleviate crowding specifically for musical stimuli. Holistic processing: No main effect or interaction reached significance. Behavioral significance of the ERP effects: ERP selectivity for musical notes was predicted by all of behavioral measures. Selectivity for notes predicted with C1, N170, and CNV, in which better music readers tend to have larger C1, N170, and CNV selectivity for notes. Crowding with extra notes predicts C1 and N170. Crowding with extra lines predicts C1, N170, and CNV. Holistic processing of music sequences among experts is correlated with both C1 and N170
Reliability	-
Validity	-
Reference	Wong, Y. K. (2010). Investigations into the role of early visual cortex in expertise reading musical notation. Retrieved from ProQuest Dissertations and Theses. (UMI No. 3442190)

Y

Title	A study of the eye-movements and eye-hand temporal relationships of successful and unsuccessful piano sight-readers while piano sight-reading
Author	Leonora Jeanne Young
Year	1971
Purpose	• To identify and examine the eye-movement patterns of successful and unsuccessful piano sight-readers, for their relationships to the quality of piano sight-reading and for their temporal relationships with the position of the successful and unsuccessful piano sight-readers' hands upon the piano keyboard. • Sub-problems: To discover what eye-movement patterns are used by skilled pianists, Are there any significant differences in the eye-moment patterns of these successful and unsuccessful piano sight-refers? If so, are there identifiable eye-movement patterns which are related to a quality of piano sight-reading? What is the temporal relationship between the successful as opposed to the unsuccessful piano sight-readers' fixation point and that which he is performing at any given movement? Is there a significant different between the successful and opposed to unsuccessful piano sight-reader's temporal relationship of the fixation point and that which he is performing at any given moment, and is there a specific relationship which is significantly related to one or the other group of piano sight-readers?
Measurement	• Pre-Test: Sight-reading test – Prospective subjects were chosen if they could play 90%+ of the information load in the difficult excerpt. Pre-test music excerpts were graded from the recordings, against a piano performance of the correct notes. Each note was given a point for its rhythmic dimension, and a point for its pitch dimension (total of 2 points). Each dot, rest, accidental was given a value of one point. • Eye-movement • Fixation response – frequency of fixation per event in the treble, bass, or total area of the 14 events • Previewing, rereading, and on chord fixations • Multiple fixations • Consecutive and nonconsecutive fixations of a single point • Frequency of frame response; period of fixation • Period of fixation off-chord • Frames without fixation spots • Time factor in consecutive and nonconsecutive fixations • Non-fixation • Eye-movement patterns • Areas fixated • Direction of movement • From-to patterns • Distinctive patterns Summary – Stationary record of moving eye-movement patterns was made, and frequency of fixation and frames and the type patterns and practices of fixation in the treble, bass, and total areas of music were tabulated for the successful as opposed to the unsuccessful piano sight-readers and analyzed for general patterns as well as group differences. In addition, frequency of fixation and frame response and the ey-hand temporal relationship about points of error were factored and tabulated to determine the changes, if any, which occurred in these types of visual behavior at points of error.
Assessments Used	-
Description	
Subjects	13 (though only 9 completed eye-movement tests) successful piano sight-readers

	(Group I) and 12 (thought only 8 completed eye-movement tests) unsuccessful sight-readers (Group II) (based on pre-tests)
Instruments Used	Piano
Stimuli	Pre-Test: Test of technical fluency - E major scale, D minor arpeggio Pre-Test: Sight-reading test – 2 musical excerpts (Nicholas Medtner and Karol Syzmanowski) for moderate and different levels, respectively. Eye-movement tests: All music examples were composed. Total of 8 examples. The example that was analyzed was D major, 4/4 time, chordal with traditional functional harmonic structure. 19 4-part chords (2 parts in each hand), and no added accidentals. RH part has 90 bits of information and the LH has 82.
Apparatus	Music Writer typewriter to produce all examples Built 'ledge' to sit on piano to hold all examples at same height *Webcor* tape recorder for each subject's performance Electric metronome used to establish tempo for each task Switching mechanism to record the moment at which each music event was played (mechanism placed in the bed under each key. As soon as the key was depressed, the switch sent electric current to oscillograph) *5-124 Consolidated Electrodynamics Corporation Recording* oscillograph to record data Modified Westgate Eye-Movement Camera, Model EMC-2 to record pattern of eye-movement (sits on helmet and photographs the movements of a white spot of light reflected off of the surface of the eye). Another camera records what is being seen. Result is a record of the eye-movement pattern superimposed on the scene of that being viewed) Counterweights to relieve weight of camera and helmet on subject Bite-bar Film
Method	Pre-Tests for prospective subjects: Test of technical fluency, Sight-reading test. For pre-tests, subjects sat at grand piano and were asked to play the scale, the arpeggio, the moderate excerpt, and the difficult excerpt. Subjects were not allowed to look at sight-reading before performing. With apparatus, subjects were seated and then calibrated before performing. Subject was told that all music examples would have D major key signature and 4/4 time signature. Tempo was set with metronome, and shut off as subject began to play. While music was put in place, subject was to depress middle C and look at upper left square until told otherwise.
Results	Only analyzed the chordal example read by subjects. Found that the eye makes several fixations in one area and then moves down to another area of the chord, then back up to another or the first area. Thus, eye zigzags up and down the chord, but also zigzags to left and right of the notes as it reads out the chords. Found that re-fixations occur almost constantly in film records of all subjects in both groups. Occasionally, some subjects read across several chords simultaneously. Subjects were able to sustain a rate of fixation as high as 40 fixations per second for short periods. Found that successful subjects made more fixations on the average than did their counterparts, and made these fixations in a ratio of approximately 2 fixations per note in the chord. Neither group consistently read both members of an interval with only one fixation. Successful subjects fixated in the bass clef more consistently. Found that successful readers spent 7.2% of total fixations previewing, whereas Group II spend 16.6% previewing. Found that Group I did more rereading than Group II. Group II used multiple fixations which jumped across one or more events and consecutive and nonconsecutive fixations of a single point. Average frequency of fixation per event was shorter for Group II. (Details of results in study beginning at page 558)
Reliability	"Several examples were re-recorded 2-3 times to establish the reliability of the

	method, and the method was found to be very reliable." (p. 180)
Validity	-
Reference	Young, L. J. (1971). A study of the eye-movements and eye-hand temporal relationships of successful and unsuccessful piano sight-readers while piano sight-reading. Unpublished doctoral thesis, Bloomington: Indiana University.

Z

Title	Evaluating new approaches to teaching of sight-reading skills to advanced pianists
Author	Katie Zhukov
Year	2014a
Purpose	• To evaluate 3 teaching approaches to improving sight-reading skills against a control in a large-scale study of advanced pianists • To investigate whether sight-reading skills in advanced pianists can be enhanced with training, and which teaching approach is most effective
Measurement	• Custom-made software for test analyses: Software analyzed MIDI files of performances and provided 4 numerical results: 2 on pitch accuracy and 2 on rhythm accuracy. Measured beat adjustment (number of times playing was interrupted, beat missed or skipped), extra notes, missing notes, RMS accuracy (average number of timing errors per correct note played). Subjects also answered background questionnaires, kept practice diaries during the training programmes, and take part in brief exit interviews. All interviews were transcribed, including the answer to the question of additional sight-reading: 0 = None, 1 = Little, 2 = Moderate, 3 = Extensive (with regard to how much additional sight-reading was undertaken during the study)
Assessments Used	-
Description	
Subjects	100 pianists; all advanced pianists (Eighth Grade and above of Australian Music Examination Board), in 4 equal groups. Subjects from 3 higher education institutions, 3 high schools and private studios in Queensland, Australia. Placed in training/control groups at random
Instruments Used	Piano
Stimuli	Pre and Post Test Pieces: Different for each programme, but of similar difficulty level. Baroque and Classical for Style Programme, Rhythmically complex for Rhythm Programme, Baroque, Classical, and 20th Century for Accompanying and Control groups. Each programme had chosen pieces at an easier level than the subjects' level Accompanying Training Programme: 2 well-known short pieces for violin/flute and piano each week (total of 20 pieces) Rhythm Training Programme: 1 or 2 short exercises each week (total of 18 exercises) involving complex rhythms. Style Training Programme: Focus on Baroque and Classical styles. Subjects played 2-3 short pieces per week) total of 24 pieces)
Apparatus	*Cubase* software for recording pre- and post-test performances Metronome for pre- and post-tests *Rolland* keyboard for pre- and post-tests Custom-made software for analysis of tests: Imported MIDI data from digitally captured performances
Method	10 weeks of training. Researcher met with each subject at least 3 times over the 10 weeks. At these meetings, the subjects played the materials from the preceding weeks and answered questions regarding rhythmic and structural analyses, and approaches to practice. Accompanying Training Programme: Subjects learned accompaniment part for each piece and then practiced once a week with a partner (one pianist playing the solo part and the other playing the accompaniment, and vice versa). Goal of this programme was to improve flow of playing, horizontal eye movement across staves, maintaining pulse, and counting. Rhythm Training Programme: Before playing each exercise, subjects had to answer a checklist, which guided their analysis of music, focusing on rhythmic aspects. Goal of this programme was to improve understanding and execution of rhythms. Style Training Programme: Subjects played short pieces, with a checklist focusing on

	analysis of harmonic and structural aspects of music. Goal was to develop differentiation between the two styles, such as understanding of structure, typical formulas, harmony and phrasing to develop pattern recognition and prediction skills within the two styles. Control Group: No additional input from researcher and continued their typical study and extra-curricular activities. All subjects were tested on 3 examples before and after the training programmes. Control group tested at beginning and end of semester. Pre- and post-tests followed strict protocol: Subject given 1 minute of perusal time before each piece. Metronome set with moderate tempo for each example.
Results	Found that: Accompanying group improved in all 4 categories in every test example: decreasing numbers of errors in beat adjustments, extra notes, missing notes, RMS accuracy from pre- to post-tests in test piece 1, 2, 3. However, this group demonstrated least improvement among all participants. Rhythm group improved in all categories in test piece 1 and 3, but in test piece 2 had an increase in missing notes and RMS accuracy at post-test Style group had slight increase in extra notes in test piece 1 and 2 at post-test, but otherwise improved in all categories Control group also improved in all categories and all test pieces at post-test. Further investigation into this improvement revealed that 8 of 25 control subjects did extensive self-motivated sight reading activities; more than subjects in all other groups. Found that training does have a positive impact on various aspects of sight-reading and that additional sight-reading activities improve post-test performance. Found that significant results for rhythm group suggest that rhythm training improves the overall flow and continuity of playing. Suggested that fewer interruptions and continuity of playing when sight-reading produce fewer extra notes, implying that rhythm training and participation in sight-reading activities has positive effect on reducing wrong notes in sight-reading. Suggested that extensive sight-reading activities can have positive effect on reducing missing notes during sight-reading, so that as player becomes accustomed to sight-reading regularly he/she omits fewer notes when playing. Also training focusing on particular styles does improve pitch accuracy in sight-reading of the same styles (as suggested in previous research: Sloboda, 1984; Thompson and Lehmann, 2004; Waters et al., 1998). Found that the 'style' group had improved pattern recognition and prediction skills, which allowed for more time to focus on rhythmic accuracy of their playing
Reliability Validity	Custom-made Software: Validity and reliability of import process verified with a series of manual checks. Software was trialed on pilot data to validate its effectiveness. Algorithm developed for the programme provided greater consistency and reliability of analysis of pre- and post-tests than manual marking of errors [utilized in previous sight-reading studies such as Banton, 1995; Betts & Cassidy, 2000; Kostka, 2000; Meinz & Hambrick, 2010; Penttinen & Huovinen, 2011]
Reference	Zhukov, K. (2014a). Evaluating new approaches to teaching of sight-reading skills to advanced pianists. *Music Education Research, 16*(1), 70-87.
See Also	Zhukov, K. (2014b). Exploring advanced piano students' approaches to sight-reading. *International Journal of Music Education, 32*(4), 487-498.

Index

A

B

C

D

M

N

P

R

S

T