26 Address correspondence to Carol M. Hayward, College of Musical Arts, Bowling Green State University, Bowling Green, OH 43403; e-mail: chaywar@bgsu.edu. Journal of Research in Music Education Volume 57 Number 1 April 2009 26-36 10.1177/0022429409332677 Relationships Among Music Sight-Reading and Technical Proficiency, Spatial Visualization, and Aural Discrimination Carol M. Hayward Joyce Eastlund Gromko Bowling Green State University, Ohio The purpose of this study was to examine predictors of music sight-reading ability. The authors hypothesized that speed and accuracy of music sight-reading would be predicted by a combination of aural pattern discrimination, spatial-temporal reasoning, and technical proficiency. Participants ($N = 70$) were wind players in concert bands at a medium-sized university in the Midwest. In a regression analysis with music sight-reading as the criterion variable, aural-spatial patterning and technical proficiency explained 51% of the variance, $F = 37.34$, $p < .0001$. These results support previous research that suggested that auditory, visual, spatial, and kinesthetic activations occur in coordination when wind players sight-read music notation. The results of the regression analysis suggested that although aural-spatial skills and technical proficiency skills were orthogonal, or separate, they both were essential to the complex task of sight-reading. Keywords: sight-reading; technical proficiency; aural perception; spatial-temporal reasoning; music literacy The skill of music sight-reading—the ability to read and play music at first sight—is highly valued in the field of music education. The inclusion of music sight-reading at state contests in secondary schools suggests that the ability to read and play with speed and accuracy is an important indicator of music achievement. Despite the value placed on music sight-reading at the secondary school level, auditions at college and university levels are often based solely on performance ability and do not include sight-reading. Although research has shown that "better sight-readers tend to be better performers" (Lehmann & McArthur, 2002, p. 142), we wanted to know whether the reverse was true, that is, whether performance ability predicted music sight-reading ability, especially when combined with other skills that research studies had shown were significant predictors of sight-reading ability.   Research has shown that music sight-reading is a complex skill. In an early study with college wind players, Elliott (1982) found that music sight-reading was predicted by rhythmic sight-reading and performance ability. In a study with high school wind players, Gromko (2004) found that music sight-reading was predicted by scores on tests of reading comprehension (of text), aural discrimination of rhythmic patterns, and scores on tests of spatial-temporal reasoning. The link to reading comprehension is reasonable because, like text, music notation is read left to right, and speed and accuracy of the visual scan is necessary for comprehension of the information. The link to spatial-temporal reasoning is reasonable because, unlike text, music notation consists of graphics that convey acoustical features of sound. The music staff serves as a frame on which pitches and rhythms are represented in vertical and horizontal space. For the fluent sight-reader, then, a visual scan across music notation will yield a mental image of the music's pitches and rhythms, as well as a memory for the experience of rendering those pitches and rhythms kinesthetically. As such, the reading of music notation is an integrated auditory, visual, spatial, and kinesthetic process. It is not surprising that several studies have shown that music instruction at a young age stimulated children's aural, visual, and kinesthetic development in ways that enhanced their performance on spatial-temporal reasoning tasks (Bilhartz, Bruhn, & Olson, 2000; Gromko & Poorman, 1998; Hetland, 2000; Rauscher et al., 1997; Spelke, 2008). In Hetland's report of a meta-analysis, she noted that the effects of music instruction were stronger for those children whose music experiences were reinforced with reading notation. Preschoolers in the study by Rauscher et al. played keyboards and referenced children's piano books. Preschoolers in Gromko and Poorman's

study sang, danced, played small xylophones, invented notations, and touched iconic notations. The preschoolers' invented notations were mnemonics that recalled their experiences of singing, dancing, moving, and playing, and as such, their notations encoded aural, visual, and kinesthetic information. Researchers also have found that music instruction at a young age enhanced children's performance on phonemic awareness tasks (Gromko, 2005; Lamb & Gregory, 1993; Saffran, Johnson, Aslin, & Newport, 1999; Wandell, Dougherty, Ben-Shachar, Deutsch, & Tsang, 2008). When children sing, dance, play, and read developmentally appropriate music notations, their experiences activate areas of the brain that process auditory, visual, spatial, and kinesthetic information. For experienced musicians, activations occur in synchrony as visual notation is processed spatially within tonal, rhythmic, and harmonic contexts (Sergent, Zuck, Terriah, & McDonald, 1992). Thus, the comprehension of music's meaning both draws on and is predicted by a network of skills and abilities that neuroscientists are finding is reflected in the brain's physiology. Based on the research that has shown links between music instruction and learning in other domains, neuroscientists in the Dana Consortium looked for neurological evidence in support of the links between the arts and areas outside the arts. In his introduction to the Dana Arts and Cognition Consortium Report, Gazzaniga (2008) stated that a neuroscientist begins with an observation that "a certain kind of brain activity works in concert with a certain kind of behavior" (p. vi). When strong correlations are found between learning in the arts and learning in areas outside the arts, the neuroscientist focuses on the brain mechanisms that undergird those relationships. In a series of studies, neuroscientists in the Dana group found several mechanisms by which links between learning in the arts and learning in areas outside the arts might be explained. Mechanisms included attention (Posner, Rothbart, Sheese, & Kieras, 2008), memory skill (Jonides, 2008), sensitivity to geometry (Spelke, 2008), reading fluency (Wandell et al., 2008), cognitive control (D'Esposito, 2008), working memory (Dunbar, 2008), and adult learning of a new language (Pettito, 2008). In other words, neuroscientists found that the physical structure of the brain reflects the motor and auditory experiences that characterize high levels of music performance over many years and these changes manifest themselves in tasks that require attention, memory, proportional reasoning, reading, and learning of languages, for example. In other studies that examined physiological differences in brain structure as a result of experiences in music, neuroscientists found that experienced musicians differed from nonmusicians in their brain physiology (Pantev et al., 1998; Schlaug, Jäncke, Huang, Staiger, & Steinmetz, 1995; Wandell et al., 2008). Professional pianists and violinists were found to have larger anterior portions of the corpus callosum, the area of the brain that contains "fibers from the motor and supplementary motor areas" (Altenmüller & Gruhn, 2002, p. 72). Several neuroscientists have sought to identify brain mechanisms that might explain links between learning in the arts and learning in areas outside the arts (Douglas & Bilkey, 2007; Gazzaniga, 2008; Ho, Cheung, & Chang, 2003). Douglas and Bilkey (2007) found that adult musicians (with high levels of sensitivity to tonal relationships) performed significantly better on mental rotation tasks than did adults with amusia (an insensitivity to tonal relationships). On the basis of their results, Douglas and Bilkey concluded that "melodic amusia is highly correlated with poor performance on a task that requires the manipulation of objects in space" (p. 919). Such a result would be hypothesized if pitch discrimination and mental rotation tasks shared a representational framework. On the basis of tests that compared amusic, nonmusician, and musician groups on their ability to perform pitch discrimination and mental rotation tasks simultaneously, Douglas and Bilkey found that amusic individuals were less susceptible to interference when performing tasks that combined spatial and tonal components. They concluded that "amusia is strongly linked to a deficit in spatial representation or processing” (p. 919). Conversely, musicians were highly susceptible to interference when performing tasks that combined spatial and tonal components, suggesting that auditory processing has a spatial component. What emerges from the research in neuroscience is an understanding that music   experiences are processed in several areas of the brain. Auditory, visual, spatial, and kinesthetic areas may all be engaged simultaneously in an integrated process that creates a shared representational framework. That music and spatial-temporal reasoning may share a representational framework was indicated by a study that Gromko (2004) conducted with high school wind players. Within her sample, she found evidence that reading comprehension (of text), rhythmic pattern discrimination, spatial-temporal reasoning, and styles of visual perception explained 48% of the variance in music sight-reading ability. In an earlier study with college-age wind players, Elliott (1982) included technical proficiency among his set of musical predictor variables. In his study, Elliott explained 88% of the variance in music sight-reading ability with a combination of rhythmic sightreading ability and performance ability. The strong correlation between rhythmic sight-reading ability and technical proficiency might be due to a common cognitive process that he identified as pattern detection (Elliott, 1982). Gromko and Elliott both measured sight-reading ability with the Watkins-Farnum Performance Scale (Watkins & Farnum, 1954), which consists of 14 melodic exercises that gradually increase in melodic and rhythmic difficulty. Elliott’s results suggest that those who recognized the recurring patterns within each exercise were predisposed to play the exercise with speed and accuracy, and they did so if they possessed the technical proficiency to render their imagined sound audible (e.g., McPherson & Gabrielsson, 2002). Gromko’s results suggest that individuals who performed well on tasks of reading comprehension, spatial-temporal reasoning, and pattern discrimination were predisposed to sight-read music with speed and accuracy. In a landmark study that investigated the effects of foot tapping and hand clapping on improvements over time in the sight-reading of rhythms, Boyle (1970) used the rhythms only from the Watkins-Farnum Performance Scale as his pretest and posttest. After 14 weeks, students in the experimental group who had learned to recognize the beat, to clap rhythmic patterns while tapping the beat, and to play rhythmic patterns while tapping the beat with their foot showed significant improvement in their ability to sight-read rhythms. The link between kinesthetic movement and reading ability is one that McPherson, Bailey, and Sinclair (1997) also found in their path analysis that included five music skills (e.g., sight-reading, playing by ear, playing from memory, performing rehearsed music, and improvising) and four environmental factors (length of study, quality of study, enriching activities, and early exposure). More capable musicians associated fingerings of learned music with recordings of the music, and when listening, they could be observed fingering along. McPherson et al. (1997) found a strong relationship between playing by ear, which involves integrated auditory and kinesthetic processing, and sight-reading ability. What emerges from the research on sight-reading is an understanding of the complexity of the sight-reading task and the possibility that high levels of coordination across auditory, visual, spatial, and kinesthetic areas might be both beneficial to and an outcome of music sight-reading.   Thus, research conducted in music education has shown that sight-reading was improved through teaching approaches that integrated auditory, visual, and kinesthetic processing and influenced by playing by ear, which integrates auditory and kinesthetic processing. Furthermore, sight-reading was predicted by rhythmic sight-reading and performance ability and, in a different study, reading comprehension (of text), aural discrimination of rhythmic patterns, spatial-temporal reasoning, and styles of visual perception. Therefore, the purpose of our study was to examine the relative strength of three predictors that require different types of processing: technical proficiency (kinesthetic/aural), aural discrimination of tonal and rhythmic patterns (aural/spatial), and spatial-temporal reasoning (visual/spatial), each of which has been shown to be related to music sight-reading ability; all of these predictors have not been included as predictors in a single study. On the basis of the research reviewed, we reasoned that if auditory, visual, spatial, and kinesthetic activations occur in coordination when musicians sight-read music, then speed and accuracy of sight-reading (our criterion variable) would be predicted by a combination of three predictor variables: aural discrimination of tonal and rhythmic patterns, scores on tests of spatial-temporal reasoning, and technical proficiency. Method Participants. A letter that described the study and requested volunteers was given to all wind players in two auditioned concert bands at a medium-sized university in the Midwest. Seventy agreed to participate, of whom 62 were music majors. Eleven of the 70 participants were graduate students; all others were undergraduates. Thirtythree were members of the top ensemble; 37 were members of the second ensemble. Thirty-nine played woodwind instruments (e.g., flute, oboe, clarinet, saxophone, and bassoon), and 31 played brass instruments (e.g., trumpet, French horn, trombone, euphonium, and tuba). Thirty-two were male; 38 were female. Procedure. Ensemble members were evaluated individually for their sight-reading ability and technical proficiency. Ensemble members were tested for their music sightreading ability using the Watkins-Farnum Performance Scale, Form A (Watkins & Farnum, 1954). Parallel forms reliability coefficients range from .87 to .94 (Boyle, 1992). As Boyle noted, “the test is meant to be scored as it is administered” (p. 260), therefore, one judge, a graduate student with experience as a band director, test administrator, and wind band performer, evaluated individual tests on-the-spot for pitch and rhythmic errors. Participants were tested for technical proficiency with a researcherdesigned test modeled after Elliott (1982). The test consisted of ascending and descending one-octave major and minor scales, ascending and descending one-octave arpeggios, and the chromatic scale. Individual testing time was approximately 25 minutes.   Participants were tested in groups for their aural discrimination of patterns with the Advanced Measures of Music Audiation (AMMA; Gordon, 1989). The AMMA measures players’ ability to judge whether a pair of music patterns is the same or different and, if different, whether the tonal or the rhythmic information changed. Reliabilities for the AMMA were reported as .85 (split halves) and .87 (retest) for undergraduate and graduate music majors (Gordon, 1989). Students marked their answers (same, rhythm, or tonal) on a paper response sheet for 40 pairs. Administration of this test took approximately 25 minutes. Students were tested in groups for spatial visualization with three subtests from the Kit of Factor-Referenced Cognitive Tests (Ekstrom, French, Harman, & Dermen, 1976). The three subtests are paper-and-pencil tests that measure “the ability to manipulate or transform the image of spatial patterns into other arrangements” (p. 173). Reliabilities for the three subtests were reported as .81, .84, and .90. The first subtest, “form board,” presents five shaded drawings of pieces, some or all of which can be put together to form a figure presented in outline form. The test takes 8 minutes. In the second subtest, “paper folding,” students see a paper folded two or three times and then punched; they choose one of five drawings that shows how the punched sheet would appear when fully opened. The test takes 3 minutes. In the third subtest, “surface development,” drawings are presented of a three-dimensional object (much like a cardboard box) next to a two-dimensional diagram of that object (much like a cardboard box laid flat). Students indicate which of the numbered edges on the twodimensional drawings correspond to the lettered edges of the three-dimensional object. The test takes 6 minutes. Students completed both the testing for aural discrimination of tonal and rhythmic patterns and the spatial visualization tests in one session, which took less than an hour. Results Descriptive statistics appear in Table 1. There were no obvious violations of the assumptions for parametric statistical tests (e.g., skewness and normalcy). To determine whether the predictor variables were collinear, we conducted a Pearson correlation. Table 2 displays the correlation matrix for the three predictor variables and the criterion variable, sight-reading. In this preliminary analysis, the Watkins-Farnum Performance Scale correlated with AMMA pattern discrimination ($r = 0.24$, $p < .05$), spatial visualization [$r = 0.24$, $p < .05$], and technical proficiency [$r = 0.70$, $p < .001$]. Among the predictor variables, the correlation between AMMA and spatial visualization was positive and significant [$r = 0.35$, $p < .01$]. Correlations between AMMA and technical proficiency and between spatial visualization and technical proficiency were nonsignificant. Due to the correlation between pattern discrimination and spatial visualization, we chose to conduct a factor analysis.   The factor analysis procedure (rotation method: Varimax; see Table 3) revealed two factors. Pattern discrimination and spatial visualization loaded on Factor 1 (0.83 and 0.81, respectively) with an eigenvalue of 1.35. We named Factor 1 "aural-spatial patterning." Technical proficiency loaded alone on Factor 2 (0.99) with an eigenvalue of 1.00. Factors were orthogonal. With factor scores as independent variables, we ran a regression of Watkins-Farnum on Factor 1 and Factor 2. Minimum tolerance for entry into the model was set at 0.01. Both factors were statistically significant. With Table 1 Table of Means and Standard Deviations (N = 70) AMMA Visualization Technical Proficiency Watkins-Farnum Mean 60.93 35.33 100.44 59.33 Standard deviation 7.66 13.53 19.01 16.11 Note: AMMA = Advanced Measures of Music Audiation; Visualization = Spatial Visualization from Kit of Factor-Referenced Cognitive Tests; Watkins-Farnum = Watkins-Farnum Performance Scale. Table 2 Correlation Matrix of Independent Variables and Watkins-Farnum Performance Scale (N = 70) 1 2 3 4 1. AMMA — 0.35** 0.07 0.24* 2. Visualization — 0.12 0.24* 3. Technical proficiency — 0.70*** 4. Watkins-Farnum — Note: AMMA = Advanced Measures of Music Audiation; Visualization = Spatial Visualization from Kit of Factor-Referenced Cognitive Tests; Watkins-Farnum = Watkins-Farnum Performance Scale. *p < .05. **p < .01. ***p < .001. Table 3 Orthogonally Rotated Factor Structure Matrix for a Two-Factor Solution Variable Loading for Factor 1-R Loading for Factor 2-R Communality TechProf 0.06 0.99 0.99 AMMA 0.83 –0.01 0.69 Visualization 0.81 0.11 0.67 Eigenvalues 1.35 1.00 % variance 5% 46% Note: TechProf = technical proficiency; AMMA = Advanced Measures of Music Audiation; Visualization = Spatial Visualization from Kit of Factor-Referenced Cognitive Tests.   the Watkins-Farnum as the criterion variable and Factor 1 (aural-spatial patterning) and Factor 2 (technical proficiency) as predictor variables, the regression analysis yielded a two-variable model explaining 51% of the variance (e.g., adjusted $R2$), $F = 37.34$, $p < .0001$. Conclusion The purpose of our study was to examine predictors of music sight-reading ability. After a review of relevant research from music education and neuroscience, we hypothesized that speed and accuracy of music sight-reading would be predicted by a combination of aural pattern discrimination, spatial-temporal reasoning, and technical proficiency. Our results supported the hypothesis. The result that aural discrimination and spatial-temporal reasoning loaded together on one factor separate from technical proficiency suggests that sight-reading entails both reading and playing. 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. In a regression analysis with music sight-reading as the criterion variable, aural-spatial patterning and technical proficiency explained 51% of the variance. Our results support previous research findings that auditory, visualspatial, and kinesthetic activations occur in coordination when wind players sightread music notation. The results of the regression analysis suggest that although aural-spatial skills and technical proficiency skills are orthogonal or separate, they are both essential to the complex task of sight-reading. In other words, technical skills alone will not assure that the performer knows the sound that he or she wants to produce, and aural-spatial skills alone will not assure that the performer can produce the sound that he or she knows the notation represents (McPherson et al., 1997). Our results support the findings of researchers in music education who reported that technical proficiency, aural pattern discrimination, and spatial-temporal reasoning predicted the speed and accuracy of music sight-reading ability in wind players (Elliott, 1982; Gromko, 2004). When aural, visual, spatial, and kinesthetic skills are applied to the task of music sight-reading, they aid the performer in mental representation and physical reproduction of the music sound represented by the notation. Implications for Music Education Our results reinforce the importance of contextualization of kinesthetic skill development in music. As Boyle (1970) showed in his early study with school-age wind band students, activities as basic as recognizing the beat with a toe tap and clapping rhythms in relation to the beat can improve students' ability to read rhythms at sight. Furthermore, as McPherson et al. (1997) found in their path analysis, time spent in playing by ear and improvisation can improve students' ear, eye, and hand   coordination and their ability to render the sound they imagine audible. As early as 1905, Jaques-Dalcroze (1921) promoted the importance of inner hearing in coordination with development of technical proficiency: Every sound method of teaching music must be based on the hearing, as much as on the emission, of sounds. If the hearing faculties of a pupil are weak, they must be developed before he undertakes the study of theory. (p. 27) Within the sequence of presentation, music educators could assist their students to connect consciously aural-spatial imagery for sound with performance on their instruments. We suggest that music educators build an image of the sound through activities such as singing and hand signing in tonal solfège, which provides a context for the pitches, and clapping, tapping, and talking through the rhythms with a number or solfège system, which provides a context for the rhythms. Singing a melody expressively while fingering is a next step in the sequence that leads ultimately to playing the music expressively. Playing melodies by ear, playing rehearsed music by memory, and improvising over harmonic progressions are additional activities that build the ear, eye, and hand coordination necessary for reading with speed and accuracy. References Altenmüller, E., & Gruhn, W. (2002). Brain mechanisms. In R. Parncutt & G. McPherson (Eds.), The science & psychology of music performance: Creative strategies for teaching and learning (pp. 63–81). New York: Oxford University Press. Bilhartz, T., Bruhn, R., & Olson, J. (2000). The effect of early music training on child cognitive development. Journal of Applied Developmental Psychology, 20, 615–636. Boyle, J. (1970). The effects of prescribed rhythmic movements on the ability to read music at sight. Journal of Research in Music Education, 18, 307–318. Boyle, J. (1992). Evaluation of music ability. In R. Colwell (Ed.), Handbook of research on music teaching and learning (pp. 247–265). New York: Schirmer Books. D'Esposito, M. (2008). Developing and implementing neuroimaging tools to determine if training in the arts impacts the brain. In M. Gazzaniga (Ed.), Learning, arts, and the brain: The Dana Consortium Report on Arts and Cognition (pp. 71–80). New York: The Dana Foundation. Douglas, K., & Bilkey, D. (2007). Amusia is associated with deficits in spatial processing. Nature Neuroscience, 10, 915–921. Dunbar, K. (2008). Arts education, the brain, and language. In M. Gazzaniga (Ed.), Learning, arts, and the brain: The Dana Consortium Report on Arts and Cognition (pp. 81–91). New York: The Dana Foundation. Ekstrom, R., French, J., Harman, H., & Dermen, D. (1976). Kit of Factor-Referenced Cognitive Tests. Princeton, NJ: Educational Testing Service. Elliott, C.

(1982). The relationships among instrumental sight-reading ability and seven selected predictor variables. Journal of Research in Music Education, 30, 5–14. Gazzaniga, M. (2008). Arts and cognition: Findings hint at relationships. In M. Gazzaniga (Ed.), Learning, arts, and the brain: The Dana Consortium Report on Arts and Cognition (pp. v–viii). New York: The Dana Foundation.   Gordon, E. (1989). Advanced Measures of Music Audiation. Chicago: GIA Publications. Gromko, J. (2004). Predictors of music sight-reading ability in high school wind players. Journal of Research in Music Education, 52, 6–15. Gromko, J. (2005). The effects of music instruction on the development of phonemic awareness in beginning readers. Journal of Research in Music Education, 53, 199–209. Gromko, J., & Poorman, A. (1998). The effect of music training on preschoolers' spatial-temporal task performance. Journal of Research in Music Education, 46, 173–181. Hetland, L. (2000). Learning to make music enhances spatial reasoning. Journal of Aesthetic Education, 34(3–4), 179–238. Ho, H., Cheung, M., & Chang, A. (2003). Music training improves verbal but not visual memory: Crosssectional and longitudinal explorations in children. Neuropsychology, 17, 439–450. Jaques-Dalcroze, E. (1921). Rhythm music & education. London: The Dalcroze Society. Jonides, J. (2008). Music skill and cognition. In M. Gazzaniga (Ed.), Learning, arts, and the brain: The Dana Consortium Report on Arts and Cognition (pp. 11–15). New York: The Dana Foundation. Lamb, S., & Gregory, A. (1993). The relationship between music and reading in beginning readers. Educational Psychology, 13, 19–27. Lehmann, A., & McArthur, V. (2002). Sight-reading. In R. Parncutt & G. McPherson (Eds.), The science & psychology of music performance: Creative strategies for teaching and learning (pp. 135–150). New York: Oxford University Press. McPherson, G., Bailey, M., & Sinclair, K. (1997). Path analysis of a model to describe the relationship among five types of music performance. Journal of Research in Music Education, 45, 103–129. McPherson, G., & Gabrielsson, A. (2002). From sound to sign. In R. Parncutt & G. McPherson (Eds.), The science & psychology of music performance: Creative strategies for teaching and learning (pp. 99–115). New York: Oxford University Press. Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L., & Hoke, M. (1998). Increased auditory cortical representation in musicians. Nature, 392, 811–813. Pettito, L. (2008). Arts education, the brain, and language. In M. Gazzaniga (Ed.), Learning, arts, and the brain: The Dana Consortium Report on Arts and Cognition (pp. 93–104). New York: The Dana Foundation. Posner, M., Rothbart, M., Sheese, B., & Kieras, J. (2008). How arts training influences cognition. In M. Gazzaniga (Ed.), Learning, arts, and the brain: The Dana Consortium Report on Arts and Cognition (pp. 1–10). New York: The Dana Foundation. Rauscher, F., Shaw, G., Levine, L., Wright, E., Dennis, W., & Newcomb, R. (1997). Music training causes long-term enhancement of preschool children's spatial-temporal reasoning. Neurological Research, 19, 2. Saffran, J., Johnson, E., Aslin, R., & Newport, E. (1999). Statistical learning of tone sequence by human infants and adults. Cognition, 70, 27–52. Schlaug, G., Jäncke, L., Huang, Y., Staiger, J., & Steinmetz, H. (1995). Increased corpus callosum size in musicians. Neuropsychologica, 33, 1047–1055. Sergent, J., Zuck, E., Terriah, S., & McDonald, B. (1992). Distributed neural network underlying music sight-reading and keyboard performance. Science, 257(5066), 106–109. Spelke, E. (2008). Effects of music instruction on developing cognitive systems at the foundations of mathematics and science. In M. Gazzaniga (Ed.), Learning, arts, and the brain: The Dana Consortium Report on Arts and Cognition (pp. 17–49). New York: The Dana Foundation. Wandell, B., Dougherty, R., Ben-Shachar, M., Deutsch, G., & Tsang, J. (2008). Training in the arts, reading, and brain imaging. In M. Gazzaniga (Ed.), Learning, arts, and the brain: The Dana Consortium Report on Arts and Cognition (pp. 51–59). New York: The Dana Foundation. Watkins, J., & Farnum, S. (1954). Watkins-Farnum Performance Scale. Milwaukee, WI: Hal Leonard.

  Carol M. Hayward is assistant professor of music education/band activities at Bowling Green State University. Her research interests include wind ensemble repertoire and pedagogy and gender issues in wind band education. Joyce Eastlund Gromko is professor of music education at Bowling Green State University. Her research interests include the development of literacy in music, music cognition, auditory processing, and transfer. Submitted June 3, 2008; accepted January 8, 2009. Downloaded from jrm.sagepub.com at PENNSYLVANIA STATE UNIV on May 9, 2016

Keyvan Yahya, Pouyan Rafiei Fard Independent Researchers

Although the Music Sight Reading process usually has been studied from the cognitive or neurological view points, but the computational learning methods like the Reinforcement Learning have not yet been used to modeling of such processes. In this paper with regards to essential properties of our specific problem, we consider the value function concept and will indicate that the optimum policy can be obtained by the method we offer without to be getting involved with computing of the complex value functions which are in most of cases inexact. Also, the algorithm we will offer here is somehow a PDE based algorithm which is associated with a stochastic optimization programming and we consider that in this case, this one is more applicable than the normative algorithms like temporal difference method.

Music Sight Reading, Reinforcement Learning, Optimization Analysis, Dopamine, Basal Ganglia

Reinforcement Learning [1] is one of the most concerned areas in the field of Machine Learning and Artificial Intelligence. In fact, many works which have been done are relevant to several methods of Decision Making, Optimization Problems [2], Dynamical Systems and application of these methods in many fields of sciences. Also, many studies focused on Neurobiological and Neuropsychological aspects of the Reinforcement Learning. Perhaps one of the most complete reviews for these issues was made by Dayan and Niv [3]. Through that paper, they presented new findings in neural RL and then they stated apparent but tractable inconsistencies and finally they found some Crucial challenges of the RL. They mentioned many works like [4] to update the formal analysis of the literature. Besides, as [3] stated, there are many works related to dopaminergic functionalities in the process of the RL in the brain. We could refer to [5] as one of the most recent works which have been done in the same field. However, there were some other works related to inspiration of the neurobiological studies of the brain in the RL that tried to develop new algorithms for some nonlinear problems according to the learning process in the brain [6]. 2 Keyvan Yahya, Pouyan Rafiei Fard Furthermore, as we mentioned before, [4] is one of the several works which have been studied the RL from a neurobiological perspective. In that work Daw and Doya referred to a hybrid approach to the function of a number of brain areas such as the Basal Ganglia and the cortex into the Reinforcement Learning. The relation between the Basal Ganglia and the other parts usually is comprehended as reinforcing of the learning task through a similar Hebbian learning process. It is expected after some trials of the learning process which able agent to do a specific task, a relation between the Basal Ganglia and some of the cortical areas is emerged. Moreover, in musical processes like sight reading learning process which is dopamine8dependant, people could figure out some of the relations of the neurotransmitters and the corresponded areas [7].

! We could assume that one of the first clues in our work was given by Zatorre [8] which offered some ideas about the importance of musical studies in the field of Neuroscience: “music provides a tool to study numerous aspects of neuroscience, from motor8skill to learning to emotion”. Also another studies such as [9] and many other studies in the field of neural substrates of music, suggest that music has biological roots, regarding to this fact that there are many neural networks in the human brain which are dedicated to processing of music. However Zatorre was convinced that musical processes are so complex8both psychological and musicological8 that make problems so challenging likewise many other researchers concluded the same issue in their studies such as [10] [11]. In fact [10] identified another answer to the question of importance of music in neural studies, which said: “music makes some unique demands on the nervous systems, an understanding of which should in turn help to reveal particular aspects of neuronal functions”. In [10], the author continued with suggestion of importance of some brain areas like ‘Cerebellum’ and ‘Basal Ganglia’, specially cerebellum for online error correction based on feedback, which would also contribute to optimization of timing in music performance. As another clue, studies like [12] [13][14] gave some ideas about the importance of auditory feedback in the process of memorizing, learning and performing music. Maybe the best relevant notion in the neuroscientific studies of music we have yet known is the one which asserts brain substrates of the real music processes are formed individually and quickly adaptive including widely distributed neuronal networks in both hemispheres [11]

These reflections and projections are among the important clues to extract the correct behavior of the learning processes. On the other hand, beside all of the accomplished researches in plasticity, it is shown that the plasticity notion satisfy in music as well as other fields like chronic pain. In music, the musicians’ brain holds the ability to better cope with the new conditions via adjusting the synaptic weights [15] [16] [17] [18]. Computational

Model of Music Sight Reading:A Reinforcement Learning Approach 3 About the relation between the RL and music in the literature, totally we can consider that this relation has only been limited to the use of the RL methods in the musical applications [19][20][21][22][23], whereas none of these studies, concentrated on music as a food of the RL. In this paper, we present some new perspectives of the reinforcement learning in music sight reading. at first, we give a short introduction about the sight reading problem and then offer a model based on the corrected aspects of the RL and finally examine the potential of our approach to be extended in other cases and conclude that with regards to this approach and the other findings, the normative theory of the RL cannot be successful in even the cases which our ancestors suggested [5]. What we really need to model sight reading is to impose a 'policy learning rate' which is a flexible entity into the computations. We show that in our special problem, the optimal learning policy is even affected by essential constraints of the problem.
"
#
$

Music Sight Reading is a multi aspect process in which many parts of the brain are involved with. This ability relates to receiving a highly complex visual input under the constraints of real time to process them while there is no opportunity for error correction [24][25]. Many experts believe that sight reading is a particular interest for professionals (piano players, conductors, orchestra players, composers) while there are stronger evidences stated sight reading is one of the five basic performance skills everyone who does music should acquire [26]. Also, specifically the important role which sight reading plays is emerged when many people are unable to read music but a vast majority is motivated to learn, as Stewart asserted [27]. To perform music sight reading process, three main sub8processes contribute and gather in the brain, namely the

and

processes according to a component skill model by [28]. The authors in [29] described sight reading as a cognitive process so that primary visual cortex plays the first role to receive the visual data then after these data can be divided into three categories in terms of how data processing will be formed in the brain. They also listed the three categories respectively as visual encoding, transcoding and production so that in visual encoding a nonverbal notation is perceived and undergoes to become abstract internal representation and finally would be ready to be produced due to playing/ singing, naming and writing. To get more details, we refer the readers to one of the best reviews which have been done by Lehmann and Kopiez [30]. %
!
&

'(

Learning of sight reading is a dopamine8dependant process. Many processes are considered following behaviors which are similar to the ones issued by the RL. For example, some people did such investigations on what does happening in the 4 Keyvan Yahya, Pouyan Rafiei Fard human brain while releasing dopamine and concluded that dopamine releasing can be held responsible for the RL [31]. They claimed that dopaminergic activities which are recorded in the Substantia Nigra and Ventral Tegmental Area (VTA) is corresponded somewhere in the brain with a kind of reward8prediction and prediction error which would be called up. In this model the prediction error is projected from the substantia nigra to the Basal Ganglia [4]. This projection is usually modeled via the Temporal Difference Algorithm. Of course, there are some limitations which were mentioned in some critical papers like [5]. The author offered some reasons to indicate that we need the precise time of impulses issued from the rewards of which the agent receives. Knowing the precise time of impulses and to where these impulses may happen is essential for the model to be succeeded. Generally, in the Reinforcement Learning, we have a state space on which decision making would be done, an action space and a policy : → which assign an action and . Here we suppose our state as the phrases extracted from musical notes. We must care that our state is not single notes. “…

[32]. The human brain is doing the learning tasks via the multiple subsystems [6]. It is strongly believed that the Basal Ganglia have been specified to instrumental conditioning via the RL [5]. But one of the most important problems in applied theories of the RL referred to dopamine releasing which is supposed to obey computational steps of the value function. But in [33] this theory which was accepted has extensively been challenged and the authors could show that just only a few dopaminergic cells projects impulses as the prediction error from the substantia nigra to the Basal Ganglia. Here we suggest our theory that is constructed based on an optimization problem rather than the value function notion. In fact, our theory would be useful when we encounter a problem which has some constraints and the learning task is separated between those constraints. For instance, the music sight reading process has this property as we will describe later. Therefore, we need an optimization based solution for this problem which cannot be overcome by the RL. In the classical algorithms to get the optimal policy, we often do some of the computations likewise Sutton suggested in [1], that is solving the Belmann Equation then after finding the state which maximize the value function as following algorithm:

= | , , (1) | , = + ∑ | , , (2) ∗ = !"#

, (3)

← 1 − '(

+ '(+ "#)

. (4) One important thing can be taken into account is that the states in musical cases are essentially related to human perception and therefore require a different analysis with

respect to the pure phenomena like frequencies or pixels. The states in musical cases have the hidden markov property because visual perception exists on them (musical notes are received as visual data in primary visual cortex). By concerning this point, we must take as a combination of a markov and hidden markov states as Mumford suggested [34]: = , + . (5) where + stands for a hidden state. By taking ,'s as the set of all parameters we have a 'Likelihood form of probability function': Pr |, = Pr | +, ,. / + |,. (6) In all of the following equations, the states 0 's are so considered that S is the state space (the groups of musical intervals like {A − E, F − B, D − G}) and A is the action space (each possible neighbor frequencies of the read note to arrive the correct interval like {9 : ;#=D# ← EA is the origin → F FG # H I}).)

Partially Observable Markov Decision Process of music sight reading. the above row is the target notes (action) the middle row is the hidden states and the bottom row is the origin intervals (state) and totally each action can be computed by / = ∑ /| +J 0 /9 +J =. Suppose too, the learning task is captured between some finite constraints denoted by K , KL . The K (rhythm) is just regulated succession of some definite pattern which is created by the movements of the first notes onsets.The KL (pitch) is supposed to be a frequency depended threshold [35]. K , KL try to affect the final result of the process and oppose the other too. Because of such an opposition, the learning policy rate denoted by M is difficult to be adjusted and obtained. What we are going to concentrate on is to giving a model which returns a proper learning policy rate. Now as we noted before, because the agent is going to move towards the definite task such that it is tended by each constraint, the neural weights never would be fixed through the learning processes and optimal policy would not be found. Our aim is doing some computations which can appear the optimal policy. The constraints are considered here to have a common intersection and thus optimal policy must belong to this common area which is subset of the Cartesian space. Suppose M is the solution of the differential equation returning the learning policy 6 Keyvan Yahya, Pouyan Rafiei Fard rate.

An important point which must be noted is that in this problem we work with stochastic variables and therefore we have to associate a stochastic analysis with optimization programming. We can write the optimization problem as following: First of All, let us substitute the prediction error term in (4) by the ‘Rescorla8Wagner surprise term’ to get: NOP =
QRS + M. . (7) Because we know from experience that when we are involved with weak reinforcers like tones or beams in so many steps the received rewards are neutral despite of the problems in which we encounter strong reinforcers such as food, sex and etc [5].)" The Reinforcement Model of Music Sight Reading Let T: → ℝ is the utility function which avoids drifting utility of the problem. Let V0 is the probability of receiving a specific state (here an interval chosen from its interval group), ∗ ∈ = { , L , … , N}.Thus the optimization problem can be written as: M ∗ : Y → ℝ in which Z[∗ (Z = −∇T9K]=^] + ^"] . (8) where is the state in which the agent resides, T is the utility function, and " is the markov process. We claim that to get an optimal policy which enables the agent to fix its synaptic weights properly, this parabolic diffusion type equation will return the policy rate. The equation (8) does have to some extent the similar attributes of the Computational Model of Music Sight Reading:A Reinforcement Learning Approach 7 famous Fokker8plank equation. When we gain the optimal rate we put it to the following equation which enables us to reach out the optimal policy: , = , + M∗ . (9) We can adopt the algorithm with the real sight reading process. When the player reads a specific note (like C), this note is transited to a new note (a new state like F) in addition a reward is received. When the state transition is occurred, the old state is weighted by this state transition: _(, = / = , ` = 1| + , … , + (10) and the prediction error between these two states is computed as: (, = _(, − a , (11) where ` is a parameter by which we mean if ` = 1 then we can ensure the state has been transited to the state . Also, + notes to the observed states till time] + 1. * +

As we mentioned above, the musical sight reading is a complex multi aspect process which has not been yet regarded to have learning property. Henceforth, now we use the method we have already described to model this process. In [7] it is shown that when someone is reading a group of musical notes from paper after receiving visual input and some preliminary processes, some activities can be detected in cortical areas. Rhythm is kept due to activities in the cerebral cortex whilst the real Pitch is perceived in the second auditory cortex as [27] suggested. The constraints can be written respectively for rhythm and pitch as the following system: b c d K], = e fgh i L j { k j − l k j } , KL], ≤ !], /440, Where K is defined as a FFT model for rhythm and constructed on the basis of the features like number of note blocks (r) and the magnitude and frequency (gh) of in STFT of underlying impulses. In KL , ! refers to the frequency of the sound the agent produces when reading note in] time and 440 is the real pitch of the ‘Tonic’ (the first and most important note of the scale on which the music ends). The first constraint is related to maintaining the real pitch of the note as far the agent is reading the sheet. The second auditory cortex serves as a servo mechanism in association with prefrontal cortex. An impulse issued from the secondary cortex to the prefrontal cortex tries to provide the notes with restoration. In fact the error, because of interfering of the constraints, cannot be predicted exactly as it is expected from (4). Another constraint as would be emerged like a vector, oppose to the pitch too. The more exact the rhythm of a phrase, the less exact the real pitch of it, as shown in Fig 3. So as it is followed some kind of the uncertainty principle which we can describe as: ΔK × ΔKL ≥ v, (12) 8 Keyvan Yahya, Pouyan Rafiei Fard where ΔK and ΔKL are the errors which can be interpreted as the difference between the real rhythm/pitch and the rhythm/pitch the agent is producing. is the prediction error and v is a constant which holds the (12).)% Oppositional behavior of the

Constraints in which the optimal policy would be emerged Thus we can rewrite our optimization analysis with regards to these constraints; the function M ∗ is represented as M∗], . The equation (8) can be considered as a free boundary problem too. Our method have an obvious advantageous respect to the other methods, that is most of the methods suffer from not to being convergent through a optimal policy series as Niv noted in [5]. Since we have a series of the finite states which are brought as input data to the visual sensory of the agent through a bounded time interval, we sort all the solutions of (8) which constructs a sequence as Γ = {M∗ , M∗ L , … , M∗ }. Γ, apparently does not has an accumulate point and thus it is not convergent. But the compactness property implies we can take M ∗ 0 , M∗ w as the upper and the lower limits of the Γ. Now it just suffices to take "
!!" over this sequence: M = sup sup Γ, (13) ∈ ∈ The equation (13) gives us a local supremum of the policy space. In addition, it is more notable from the convergency view point. After putting M in the (9) we get the optimal policy which is identical with ∗ obtained by (3) and also corresponded with
∗ as optimal value function can be gained also from (7). Computational Model of Music Sight Reading:A Reinforcement Learning Approach 9 ,

It is proposed that the method we offered above can be generalized to the other applications of the RL theory. Obviously many aspects of the learning problem are open to the prospective researches. Even the classical challenge on exploration/exploitation can be discussed as an optimization problem. To get an exact model which can inhibit the limitations of the RL, all of the crucial topics can be revised with regards to what we offered here. In other way, we must take the dynamical property of learning problems into account and to do this we may need advanced mathematical tools. Constructing the policy space can be done in various forms and the constraints which will really emerge are not just like the ones we offered. They can be strengthening to other kinds of constraints. By far and large, associating the learning problem with constraints and doing optimization analysis may lead us to get over some of the problems we encountered in theory and practice. -

1. Sutton, R.S., Barto, A.G.: Reinforcement Learning: An Introduction. MIT Press, Cambridge, MA (1999) 2. Montague, P.R.: Why Choose This Book?: How we Make Decisions. Dutton Adult. (2006) 3. Dayan, P., Niv, Y.: Reinforcement Learning: The Good, The Bad, The Ugly. Current Opinion in Neurobiolog.18, 185–196 (2008) 4. Daw, N.D., Doya, K.: The computational neurobiology of learning and reward. Current Opinion in Neurobiology. 16, 199–204 (2006) 5. Niv, Y.: Reinforcement Learning in the brain. Journal of Mathematical Psychology. 53(3), 139–154 (2009) 6. Rivest, F., Bengio, Y., Kalaska, J.: Brain inspired reinforcement learning. In Saul, L. K., Weiss, Y., Bottou, L. (eds.) Advances in neural information processing systems, vol. 17, pp. 1129–1136. MIT Press, Cambridge, MA (2005) 7. Spruit, B.:Music and the Brain. In: Essay Cognitive Neuroscience (2008) 8. Zatorre, R.: Music, the food of neuroscience? Nature. 434, 312–315 (2005) 9. Peretz, I.: Brain Specialization for Music: New Evidence form Congenital Amusia. In Peretz, I., Zatorre, R.(eds.) The cognitive neuroscience of music, pp. 247–268. Oxford University Press, New York (2003) 10. Zatorre, R.J., Chen, J.L., Penhune, V.B.: When the brain plays music. Auditory–motor interactions in music perception and production. Nature Reviews Neuroscience. 8, 547–558 (2007) 11. Altenmüller, E.O.: How Many Music Centers Are in the Brain? Annals of the New York Academy of Sciences. 930 273–280 (2001) 12. Finney, S.A., Palmer, C.: Auditory feedback and memory for music performance: Sound evidence for an encoding effect. Memory & Cognition. 31 (1), 51–64 (2003) 13. Pfordresher, P.Q.: Auditory Feedback in Music Performance: The Role of Melodic Structure and Musical Skill. Journal of Experimental Psychology. 31(6), 1331–1345 (2005) 14. Pfordresher, P.Q.: Auditory Feedback in Music Performance: The Role of Transition8Based Similarity. Journal of Experimental Psychology. 34(3), 708–725 (2008) 10 Keyvan Yahya, Pouyan Rafiei Fard

15. Rauschecker, J.P.: Functional Organization and Plasticity of Auditory Cortex. In Peretz, I., Zatorre, R (eds.) The cognitive neuroscience of music, pp. 357–365. Oxford University Press, New York (2003) 16. Schlaug, G.: The Brain of Musicians. In Peretz, I., Zatorre, R. (eds.) The cognitive neuroscience of music , pp. 366–381. Oxford University Press, New York (2003) 17. Pantev, C., Engelien, A., Candia, V., Elbert, T: Brain Representational Cortex in Musicians. In Peretz, I., Zatorre, R. (eds.) The cognitive neuroscience of music , pp. 382–395. Oxford University Press, New York (2003) 18. Pacua8Leone, A.: The Brain That Makes Music and is Changed by it. In Peretz, I., Zatorre, R. (eds.) The cognitive neuroscience of music , pp. 396–409. Oxford University Press, New York (2003) 19. Collins, N.: Reinforcement Learning for Live Musical Agents. In: Proceedings of ICMC2008, International Computer Music Conference, Belfast (2008) 20. Collins, N.: The Potential of Reinforcement Learning for Live Musical Agents. In: Proceedings of MML2008, ICML, Helsinki (2008) 21. Franklin, J.A., Manfredi, V. U.: Nonlinear credit assignment for musical sequences. In Second international workshop on intelligent systems design and application, pp. 245–250 (2002) 22. Fiebrink, R., Cook, P., Smallwood, S., Trueman, D., Wang, G.: Laptop Orchestras and Machine Learning in Real8time Music Performance. In: ACM CHI 2009, Computational Creativity Support Workshop, Boston (2009) 23. Millard, M., Tizhoosh, H.R.: Tuning pianos using reinforcement learning. Journal of Applied Acoustics. 68, 576–593 (2007) 24. Kopiez, R., Lee, J.I.: Towards a dynamic model of skills involved in sight reading music. Music Education Research. 8(1), 97–120 (2006) 25. Kopiez, R., Lee, J.I.: Towards a general model of skills involved in sight reading music. Music Education Research. 10(1), 41–62 (2008) 26. McPherson, G.E.: Five aspects of musical performance and their correlates. Bulletin of the Council for Research in Music Education, 127, 115–121 (1995) 27. Stewart, L.: A Neurocognitive Approach to Music Reading. Annals of the New York Academy of Sciences. 1060 377–386 (2005) 28. Waters, A. J., Townsend, E., Underwood, G.: Expertise in musical sight8reading: a study of Pianists. British Journal of Psychology. 89(1), 123–149 (1998) 29. Schön, D., Anton, J.L., Roth, M., Besson, M.: An fMRI study of music sight8reading. NeuroReport. 13(17), 2285–2289 (2002) 30. Lehmann, A.C., Kopiez, R.: Sight8reading. In Hallam, S., Cross, I., Thaut, M. (eds.) The Oxford handbook of music psychology, pp. 344–351. Oxford University Press, Oxford (2009) 31. Schultz, W.:Predictive reward signal of dopamine neurons. Journal of Neurophysiology. 80, 18 27 (1998) 32. Welford, A.T.: Skilled performance: perceptual and motor skills. Scott Foresman Comp Glenview Ill. 1–200 (1976) 33. Redgrave, P., Prescott, T.J., Gurney, K.: Is the short8latency dopamine response too short to signal reward error? Trends neurosci. 22:4, 146–151 (1999) 34. Mumford, D.: Pattern Theory: the Mathematics of Perception. In Proceedings of the International Congress of Mathematicians: Vol. III (2002) 35. Holzapfel, A., Stylianou,Y.: A scale transform based method for rhythmic similarity of music. In Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (2009)

Abstract

Deliberate practice—that is, engagement in activities specifically designed to improve performance in a domain—is strongly predictive of performance in domains such as music and sports. It has even been suggested that deliberate practice is sufficient to account for expert performance. Less clear is whether basic abilities, such as working memory capacity

(WMC), add to the prediction of expert performance, above and beyond deliberate practice. In evaluating participants having a wide range of piano-playing skill (novice to expert), we found that deliberate practice accounted for nearly half of the total variance in piano sight-reading performance. However, there was an incremental positive effect of WMC, and there was no evidence that deliberate practice reduced this effect. Evidence indicates that WMC is highly general, stable, and heritable, and thus our results call into question the view that expert performance is solely a reflection of deliberate practice.

Psychology of Music 1 –16 DOI: 10.1177/0305735615585398 pom.sagepub.com The art of sight-reading: Influence of practice, playing tempo, complexity and cognitive skills on the eye–hand span in pianists Stephanie Rosemann1 , Eckart Altenmüller2 and Manfred Fahle1 Abstract Sight-reading is a skill required by musicians when they perform sheet music unknown to them. It demands sequential anticipatory eye fixation of notes immediately followed by motor execution. The distance between eye (fixation of a note) and hand position (tapping the corresponding key) is called eye-hand span (EHS). The aim of our study was to investigate the influence of practice, playing tempo and complexity of the music on the size of the EHS, as well as its relation to performance and cognitive skills (shape recognition, working memory, and mental speed). We used a sight-reading paradigm where nine pianists accompanied a pre-recorded flute voice, which also served as a timekeeper. After a practice phase, a second measurement of the EHS with same tempo and a third and fourth measurement with a different playing tempo followed. We found that the practice phase only slightly affected the EHS but that the EHS significantly changed according to playing tempo and complexity of the music. Furthermore the EHS correlated with quality of performance after practice and mental speed skills. Hence we conclude that the EHS seems to be characteristic for each musician, is developed over years of practice and is relatively independent of a short practice phase. Keywords cognition, complexity, eye movements, practice, sight-reading, tempo Sight-reading (SR) a piece of music is a skill which is required by almost all musicians (Wurtz, Mueri, & Wiesendanger, 2009). It is the unrehearsed performance of music and is needed when a musician plays a composition hitherto unknown to him/her. In most studies investigating the 1Department of Human-Neurobiology, University of Bremen, Germany 2Institute for Music Physiology and Musicians' Medicine, Hanover University of Music, Drama, and Media, Germany Corresponding author: Stephanie Rosemann, Department of Human-Neurobiology, University of Bremen, Hochschulring 18, D-28359 Bremen, Germany. Email: strosemann@uni-bremen.de 585398POM0010.1177/0305735615585398Psychology of MusicRosemann et al. research-article2015 Article   SR abilities of musicians, an influence of elementary cognitive skills such as memory capacity was found, as well as practice-dependent skills like cumulative practice time (Gruber, Jansen, Marienhagen, & Altenmüller, 2010; Kopiez & Lee, 2008; Kopiez, Weihs, Ligges, & Lee, 2006; Lee, 2003). During SR, the musician reads the notes and has to translate this visual-symbolic representation into appropriate movement patterns.

This implies particular requirements on the musician in terms of processing a highly complex visual input within a certain time frame and simultaneously processing perceptual, cognitive and motor demands (Kopiez & Lee, 2008). While sight-reading, the musician’s eyes constantly scan the notes slightly ahead of the notes being played. This distance between the notes currently played (hand position) and those fixated upon (eye position) is called the eye–hand span (EHS) and can be denoted by an anticipatory time span (length of time between the fixation and performance of a note; in s or ms) or by notes (number of notes between hand and eye position). This anticipation in reading the notes is needed to recognize, decode, process and store the sheet music as well as to plan and execute the motor behaviour. The EHS seems to have an optimum duration, at least for a particular piece and a certain playing tempo (Furneaux & Land, 1999). Thus, it is possible to address the question of how much information can be transformed from a visual domain into motor output within a certain time frame (Lehmann, 2005). Reading music is a kind of visual pattern recognition because musicians do not separately process note-by-note, but look at groups of notes, an ability referred to as ‘chunking’ (Kinsler & Carpenter, 1995; Penttinen & Huovinen, 2011; Waters, Underwood, & Findlay, 1997; Wolf, 1976). In SR literature, positive correlations between SR performance and pattern recognition skills (Waters, Townsend, & Underwood, 1998), working memory capacity (Lee, 2003; Meinz & Hambrick, 2010) and mental speed (Kopiez & Lee, 2008; Lee, 2003; Meinz & Hambrick, 2010) were identified. Additionally the relationship between different factors influencing SR achievement for increasing SR complexity has been explored (Kopiez & Lee, 2006). These authors showed that with increasing task difficulty, mental speed and expertise were the important predictors for SR performance, but that at the highest complexity level only mental and psychomotor speed became the dominant predictors. However, it was not determined how cognitive skills are related to the duration of the EHS. In the last 20 years, only a few studies have dealt with the EHS of musicians during SR (Furneaux & Land, 1999; Kinsler & Carpenter, 1995; Truitt, Clifton, Pollatsek, & Rayner, 1997; Wurtz et al., 2009). These studies used eye-movement transducers and eye-tracking devices. Earlier research on the EHS included switching off the light or the music slide and counting the number of musical events the musicians were still able to play (Rayner & Pollatsek, 1997). All of them found different lengths of the EHS and used different measures of quantification for the EHS (Rayner & Pollatsek, 1997: 1.5–3.1 notes; Sloboda, 1974: 3.8–6.8 notes; Furneaux & Land, 1999; Wurtz et al., 2009: 1 s). Truitt et al. (1997) used the ‘moving window’ technique to measure the perceptual span – the size of the effective visual field – and to combine these results with findings for the EHS. The perceptual span was between 2 and 4 beats and at most a little over a single measure and the mean EHS averaged somewhat over one and ranged from the eye positions being 2 beats behind up to 12 beats ahead of the hands. Gilman & Underwood (2003) used a similar approach to investigate the perceptual span and EHS for pianists. They found that expertise did not have an influence on the perceptual span but on the size of the EHS. Furthermore, it was shown that the EHS is affected by expertise (Truitt et al., 1997) and skill level (Furneaux & Land, 1999; Sloboda, 1974; Truitt et al., 1997). It was demonstrated that the EHS is influenced by playing tempo (Furneaux & Land, 1999) and by complexity and structure of the music itself (Rayner & Pollatsek, 1997; Wurtz et al., 2009) too. Hence, it is assumed that the EHS can be altered by the general conditions of the experiment and is determined by the skill and expertise of the musician. Besides the different computation methods for the EHS, the previous studies differed in terms of experimental procedure (metronome vs. own tempo) and type of music (simple melodies vs. dual-staved music). Therefore, it is difficult to draw firm conclusions, especially when the playing tempo was not fixed for all participants within the same experiment, although it was shown that the playing tempo affects the duration of the

EHS. As for the influence of complexity on the EHS (Wurtz et al., 2009), the EHS did not change in latency but was correlated with the playing tempo. The influence of complexity on the EHS for a fixed playing tempo remains to be investigated. The impact of the skill level on the EHS was also demonstrated and a larger EHS was found for skilled players compared to less skilled players (Furneaux & Land, 1999; Sloboda, 1974; Truitt et al., 1997), but the different groups varied in terms of expertise and playing experience. Hence, the question arises whether the EHS only correlated with the expertise and playing experience of the musicians and not necessarily with skill. The influence of a short practice phase (rehearsal) on the size of the EHS has not previously been addressed, but an increase in the perceptual span due to rehearsal has been identified (Burman & Booth, 2009). The perceptual span increased from 4–5 notes to a span of 11 notes as a result of training. Up to now, investigators used different techniques to measure the EHS and obtained mixed results regarding the influence of expertise, skill, tempo and structure of the music on the EHS, each study dealing with only one or two possible factors. Cognitive skills play an important role for SR achievement but have not yet been considered as determining factors for the EHS. The effect of a short training phase (rehearsal) on the EHS has not received any attention so far, although its influence on the perceptual span has been demonstrated: experts revealed a larger EHS than novices possibly due to intensive training. The present study aims to investigate different factors that might be linked to the size of the EHS of pianists by recording the eye-movements of expert pianists during a SR task, after rehearsal and for different tempi. We applied an 'ecological task'-approach by mimicking a piano-accompaniment situation with a pre-recorded flute voice and utilized a novel approach (technology) to precisely determine eye-movements in the millisecond range. With the use of standardized neuropsychological tests and a questionnaire, the influence of cognitive skills and expertise (defined by quantitative measures of, for example, overall life practice time and qualitative measures as marked grading of performance) was studied. Playing tempo was fixed and determined by the pre-recorded flute voice. The EHS was calculated as the number of beats between current eye and hand position and as the latency between fixating and playing a note. With this study we will be able to draw conclusions about many different factors that might influence the size of the EHS. No other study has included so many factors in their experiment before; previous research has only concentrated on one or two of them. Furthermore, we raise the issue of a possible rehearsal effect on the size of the EHS which has not been considered before by any other study. We assumed that a larger EHS would be associated with rehearsal, a slower playing tempo, and reduced complexity of the music. In addition, with increased EHS, we expected a better performance and a higher score in cognitive tests, especially in working memory, mental speed and shape recognition. To sum up, we believe that the EHS depends on general aspects of the piece of music to be played (structure, difficulty, tempo, time spent on practicing it) and on characteristics of the person who plays it (skill, hours spent on practicing the instrument, cognitive skills).   Method Participants Participants were informed about the aim and procedure of the experiment and had to sign a written consent form according to the Declaration of Helsinki. All of them were paid for participation in the experiment and were free to withdraw from the study at any time. Participants (N = 9) were Piano Students (four female, five male) from the Hanover University of Music, Drama and Media. The students had piano as their major subject (skill level was supposed to be high; no novices or intermediate piano players were included). Two participants had already completed their degrees, the other students had accomplished 6 (± 2.5) semesters of their studies. The mean age was 25.11 (± 6.47) years. Eight participants were right-handed, one was lefthanded and all of them had normal or corrected-to-normal vision. Eye-tracking required the students to have light-coloured eyes because the eye-tracking program used the boundary between light iris and dark pupil to detect the pupil and compute the fixation position. Equipment and materials All participants performed the accompaniment score ‘Adagio ma non tanto’ of the Flute Sonata in E minor (BWV 1034) composed by Johann Sebastian Bach (1685–1750). This piece was unknown to all of them. The sheet music consisted of a melody (flute) and an accompanying piano voice presented on two DIN A4 pages and comprising 30 measures in 12 rows (six on every page). The flute voice was recorded on tape by the second author, using a microphone (not produced by a Musical Instrument Digital Interface). During the SR task, the pianists had to accompany the flute voice, which served as a time-keeper to avoid timing differences between participants. Prior to the first flute tone, a metronome indicated the tempo of the piece during one bar. The eye movements of the pianists were recorded by an eye tracking system. The music performance was recorded via Musical Instrument Digital Interface (MIDI) which directed the input of the piano to an external computer. The experimental set-up is shown in Figure 1. The chin rest, the eye tracking camera and the light-emitting diodes (LED) of the eye tracking system were fixed at the table where the piano (Kawai, Professional Stage Piano MP 9000) was placed. The participants had to keep their head on the chin rest (42 cm height from the table) to which the camera was attached. The distance between eyes and camera was 20 cm. The LED lights were positioned on the piano so that the distance from the eyes to the infrared light was 43 cm and the distance between both lights was 30 cm. The music stand with the sheet music was placed behind the piano at 70 cm distance from the eyes, and the width of the sheet music (both pages side by side) was 39.5 cm. The height of the table was 68 cm and the height of the camera from the table was 40 cm. Although the piano was placed on the table below the eye tracking camera and the chin rest, there was enough space for the hands to play the piano. Software developed in-house served to track the eye movements of the pianists. A computer received the signal from the camera, which was attached to the chin rest and used the corneal Purkinje reflex of the eyes, which were illuminated by the LED lights relative to pupil position, to calculate the target position the participant was fixating. In this experiment, the left eye of each participant was used for eye tracking. Procedure In the first trial, participants were required to sight-read the music as soon as the flute voice was replayed. For this purpose, four quaver notes were counted in advance and the flute voice   started after the fifth quaver because there was a quaver rest on the first beat. No prior study of the piece of music was allowed, to minimize the effect of prior exposure. After the first measurement, the pianists were allowed to rehearse the piece of music for 30 minutes. Thereafter, the participants had to play the piece in three different tempi. First, all of them played it in the original tempo (second trial). In the third and fourth trial, the participants had to play the piece 20% faster or else 20% slower than the original one. Half of them started with the fast tempo and the other half started with the slow tempo. During all four trials, the music and eye movements were recorded. Throughout the entire experiment, all participants were able to hear what they were playing. Assessment of performance and cognition We assessed the participants’ performance for the measurement at sight and after rehearsal to compare the performance of both measurements with the same tempo. A German grading system served as a measure for the piano playing performance (so-called skill level). The participants were graded by the second author using blind marking and possible grades ranged from 1 to 6 (1 equals A, 6 equals F). Grading took into account the number of mistakes, number of omitted notes and correct phrasing. To evaluate the sight-reading and piano-playing expertise, the participants were interviewed and had to fill in a questionnaire to keep a log about their musical activities in order to calculate the accumulated hours of practice and experience (similar to retrospective interviews used by Kopiez et al., 2006). This record was divided into three categories of activities (Practice, Lessons, and Sight-reading), each consisting of two sub-categories. The sub-categories were ‘Main

Instrument' and 'Minor Instrument' for 'Practice' and 'Sight reading.' The category 'Lessons' was divided into 'taking lessons' and Figure 1. Experimental set-up. Kawai piano placed on the table where the eye tracking system was attached.   'giving lessons.' Participants were asked to fill in how many hours per week they had spent on the different categories at four different age ranges: up to 10 years old, 10–15 years old, 15–18 years old and 18 years old–today. These different age phases were used to keep track of the practice behaviour of the participants during different phases of their musical education and to calculate the overall life practice time for each participant. We ended up with seven variables as measures of expertise in relation to past experience: 1) number of concerts, 2) overall practice time for the piano, 3) overall practice time for the minor instrument (if applicable), number of 4) taken and 5) given lessons (separately), hours spent on practicing SR on 6) piano and on 7) the minor instrument (if applicable). Cognitive performance was assessed with a selection of neuropsychological tests which included tests of mental and processing speed, short term and working memory as well as shape recognition skills. A test for speed of attention, mental flexibility, visual search and motor function is the Trail Making Test (TMT; Reitan & Wolfson, 1985). It is a paper and pencil test where the participants were required to connect ascending numbers (part A) or alternating numbers and letters (part B) as fast as possible. The participants completed both parts of the test. The time the participants needed to complete each part was measured using a stop watch by the experimenter. This processing time (in seconds) was used for evaluation. A second test for assessing sustained attention and visual scanning ability is the D2 Concentration Endurance Test (D2; Brickenkamp, 1994). It is also a paper and pencil task, which consists of 14 rows with 47 characters per row. The characters are 'd' and 'p' each with one to four vertical dashes. Targets are 'd' characters with two dashes and participants were required to cross out targets and leave non-targets untagged with a time constraint of 20 s for each row. The overall number of crossed out characters, omissions and errors was evaluated to measure the quality of performance for each subject. Simple reaction time tasks were used to measure visual and auditory reaction times (self-developed). Both tests consisted of 20 trials, and for data analysis, the median reaction time was used. For the measurement of the visual short term memory, the Visual Pattern Test (VPT) was used (Della Sala, Gray, Baddeley & Wilson, 1997). The participants were required to memorize 42 matrix patterns of black and white squares with a presentation time of 3 s each. The complexity of the patterns increased from a 2 × 2 matrix to a 6 × 5 matrix. After the presentation of each pattern the participant had to draw it onto the pre-assembled scoring sheet. For scoring, the number of correct squares for each pattern was used (e.g. for the first pattern, there are a maximum of 4 correct squares, and for the last one, a maximum of 30). The digit span task, taken from the 'Hamburg-Wechsler-Intelligenztest für Erwachsene' (HAWIE), is a measure of the short term and working memory (Tewes, 1994). Participants had to recall digits of increasing length in chronological (part A) or reversed order (part B). The number of recalled digit spans and the according age-corrected value was used for scoring. To test the ability to recognize shapes, two subtests (10, 11) of the 'Leistungsprüfsystem' (LPS; Horn, 1983) were conducted in the experiment. In subtest 10, the participants were required to recognize some specified symbols (T, U, L, triangle, rhombus) in a figure (processing time: 3 min). In subtest 11, the participants had to recognize fragmented figures and mark the beginning letter of that figure. Below each figure, the target letter (initial letter of the figure) and four distracter letters were displayed (processing time: 1 min). The number of correct responses was evaluated and the respective age-corrected t value was used for scoring. Computation of the EHS and statistical analysis In order to calculate the EHS, the eye tracking program was modified to simultaneously record the eye movements and to receive the MIDI input of the played notes. Temporal information about eye movements and onsets of note performance were directly acquired and plotted on the sheet music on a monitor by Matlab-based programs (Figure 2). With this visualization, the computation of the EHS was possible as both temporal information about eye movements and note performance as well as information about the sheet music itself was available. The EHS was calculated a) as latency (in ms) between fixating and playing a note, and b) as number of beats between current eye and hand position (see the supplementary material for more details). The use of the number of beats instead of the number of notes was adequate because it serves as a constant value throughout the whole music while the number of notes varies enormously across measures. To reduce the amount of data to a manageable proportion, the EHS was computed for the first and third beat of every measure (except: first beat of measure one skipped for latency computation and first beat of measure 30 skipped for beat computation). This was done for all four trials (at sight, after rehearsal, slow tempo, fast tempo) for all 30 measures, hence a maximum of 58 data points resulted for one measurement of one type of EHS. Measures which were rated to be difficult (measures 5, 10–13, 24–26) or easy (measures 1–4, 22, 23, 29) underwent a detailed EHS computation of every beat of the respective measures in order to evaluate the EHS in terms of different levels of difficulty. The rating for difficult and easy measures was done by EA. Before the statistical analysis, data were interpolated during short periods of signal loss of the eye tracking program or where pianists made mistakes in playing. All measures for which no EHS value could be calculated for at least four participants were excluded from analysis (9 data points for measurement at sight; 4 data points for measurement after rehearsal; 2 data points for measurement with fast tempo; 6 data points for measurement with slow tempo). The other empty cells in the data pool were filled with the mean EHS of the other participants of the respective measure. So far, the EHS computation and data correction were performed separately for each trial. Next was the comparison of the EHS for different trials (e.g. to test the first hypothesis, the EHS values for the first and second trial were compared with each other). Thus, each group of trials (at sight/after rehearsal comparison, original/slow/fast tempo comparison) was compared, so each group had to contain exactly the same data points to exclude effects following from the comparison of different measures of each participant. For the statistical analysis of each

Figure 2. Visualization of eye tracking and MIDI data. Eye tracking and MIDI data (one participant and one measurement for row 7 of sheet music shown) have been plotted on the sheet music using Matlab to calculate the EHS in latency and in number of beats. The flute voice is at the top and both hands of the piano voice on the bottom two lines. Dotted line: Eye tracking data; Letters: played MIDI notes; Dark cursor left: selects eye tracking data; Light cursor right: selects MIDI data.

comparison, only those data points that could be calculated in the respective trial of each group were included. The trials ‘after rehearsal’ and ‘original tempo’ refer to the same trial but for comparison with either the trial at sight or the different playing tempi, different measures were included in the statistical analysis. Statistical analysis was performed by using IBM SPSS Statistics 20. To identify the influence of rehearsal, tempo changes and structural difficulty, repeated measures Analysis of Variance (ANOVA) and post hoc paired t tests were conducted. The relationships between cognitive skills, performance, expertise and the EHS were analysed with a non-parametric rank correlation analysis (Spearman-Rho, 2-tailed). A non-parametric rank correlation test was required because many of the variables were not interval scaled (e.g. grades, number of concerts).

Results

The aim of the study was to identify influences of rehearsal, temporal changes and different levels of structural difficulties on the duration of the EHS of pianists performing a SR task. Furthermore the relationships between performance, expertise and cognitive skills and the size of the EHS were investigated. To achieve this aim, repeated measures Analysis of Variance (ANOVA), post hoc paired t tests and a correlation analysis (Spearman-Rho, 2-tailed) were conducted. Prerequisites for ANOVA calculation were fulfilled (normal distribution and homogeneity of variances). In the following, the mean values for the different trials over all participants are discussed and displayed (see supplementary Tables S1 and S2 for mean EHS values for each participant). Results of the EHS for measurements at sight and after rehearsal A paired-sample t test (one-sided) was conducted to compare the values of the EHS in latency and in number of beats for the measurement at sight and after rehearsal (Figure 3). There was a trend in latency difference between the duration of the EHS at sight (1252 ms, SD = 135) and after the rehearsal phase (1326 ms, SD = 141), $t(8) = -1.618$, $p = .072$. For the number of Figure 3. Average EHS results for the measurements at sight and after practice. Data (mean ± standard error) are shown for measurement at sight (N = 9) and measurement after practice (N = 9). A: EHS in latency (ms). B: EHS in number of beats.   beats, no significant difference between the values at sight (0.46 beats, SD = 0.12) and after rehearsal (0.47 beats, SD = 0.12) was obtained either, $t(8) = -.395$, $p = .352$. Half a beat corresponds to approximately 4.5 notes or 1.5 chords. Three different types of EHS changes after rehearsal were present for the number of beats (Figure S2): decreased EHS, increased EHS, and no change. An analysis was conducted for subgroups of three participants each (with decreased versus with increased EHS; participants with negligible change were omitted from this analysis). Unpaired t tests showed that both groups did not differ before rehearsal ($p = .356$) but differed significantly after rehearsal ($p = .006$). Within these subgroups the change of the EHS before versus after rehearsal did not yield significance for the group with increased EHS ($p = .104$) but a trend was present for the group with decreased EHS ($p = .073$) as computed with two paired t tests. Results of the EHS for different tempi For the EHS in latency, the mean value across all participants with original tempo was 1342 ms (SD = 150), for the fast tempo 1143 ms (SD = 135), and with slow tempo 1475 ms (SD = 184). The EHS in number of beats for the original tempo was 0.42 beats (SD = 0.12), for the fast tempo 0.47 beats (SD = 0.15), and for the slow tempo the number of beats was 0.29 (SD = 0.11). The results can be seen in Figure 4. A one-factor ANOVA for repeated measurements with three levels (original, slow, and fast tempo) was calculated for the values of the EHS in latency and showed a significant main effect of tempo, $F(2, 16) = 33.017$, $p < .001$. Post-hoc paired comparisons indicated that all values differed significantly from each other. The EHS of the fast tempo was significantly smaller than that of the original tempo ($p < .001$) and significantly smaller than the EHS of the slow tempo ($p < .001$). Furthermore, the EHS of the slow tempo was significantly higher than that of the original tempo ($p < .01$). Another one-factor ANOVA for repeated measurements was performed for the EHS values in number of beats and showed a significant main effect of tempo, $F(2, 16) = 17.945$, $p < .001$. The post-hoc paired comparisons showed that the EHS of the slow Figure 4. Average EHS results for the measurements with original, fast and slow tempo. Data (mean ± standard error) are shown for measurement with original tempo (N = 9), measurement with fast tempo (N = 9) and measurement with slow tempo (N = 9). A: EHS in latency (ms). B: EHS in number of beats.   tempo significantly differed from the other two tempi but that the original and fast tempo values did not differ significantly from each other ($p > .05$). The EHS in beats of the slow tempo was significantly smaller than the EHS of the fast tempo ($p < .01$) and significantly smaller than the EHS of the original tempo ($p < .01$). Effect size measures for the effect of temporal change on the EHS in latency ($\eta 2 = 0.846$) and number of beats ($\eta 2 = 0.834$) showed a large effect. Interestingly, no proportional change of the EHS according to the change of the playing tempo (20%) took place. The EHS for the fast tempo decreased by 14.8% and the EHS for the slow tempo increased by 10%, so there was indeed a change of the EHS due to playing tempo but it was not proportional to the change of the playing tempo. Results of the EHS for different complexities Measures which were identified as difficult (measures 5, 10–13, 24–26) or else easy (measures 1–4, 22, 23, 29) underwent a detailed EHS computation in order to evaluate the EHS in terms of different levels of difficulty. To compare values of different levels of difficulty and complexity, a paired-sample t test was conducted (Figure 5). There was no significant difference of latency between the size of the EHS for difficult measures (1268 ms, SD = 149) versus easy measures (1320 ms, SD = 260), $t(8) = -1.069$, $p = .316$. For the number of beats, a significant effect was obtained, $t(8) = -4.015$, $p = .004$, with the value of the EHS for the difficult measures (0.35 beats, SD = 0.16) being significantly smaller than the value for the easy measures (0.51 beats, SD = 0.23). SR and cognitive performance To analyse the relationship between the size of the EHS, the SR and cognitive performance, as well as the musical expertise, we performed a correlation analysis (Spearman-Rho, 2-tailed). Corrections for multiple comparisons were not applied because of our complex study design and the small group size. Figure 5. Average EHS results for easy and difficult measures. Data (mean ± standard error) are shown for difficult measures (N = 9) and easy measures (N = 9). A: EHS in latency (ms). B: EHS in number of beats.   Table 1 shows the results of the Spearman-Rho correlation while Tables S3 and S4 contain data for individual participants regarding their expertise and cognitive assessment. Participants started playing the piano on average at an age of 5.7 (± 2.1) years. They had a cumulated life practice time of 14,292 hours and had lessons of 954 hours. Two significant correlations emerged, namely a) a positive correlation between the results of the D2 Concentration Endurance Test and the EHS in number of beats and b) a negative correlation between the results of the LPS 11 and the EHS in ms. For the relationship between the EHS in beats and all marked signs of the D2 test, the correlation value was 0.85 ($p = .004$). The correlation between the EHS in ms and the LPS 11 was negative: −0.707 ($p = .033$). Hence, a large EHS was associated with a low score in the LPS subtest 11. There was a trend for the correlation between the EHS in beats and the results of the TMT part A (−0.586; $p = .097$). This correlation was negative, as the score was determined by the Table 1. Spearman-Rho Correlation Coefficients between the EHS, SR and cognitive performance. EHS [ms] EHS [beats] RT_Visual −0.400 −0.267 RT_Auditory −0.033 −0.192 TMT_A −0.209 −0.586T TMT_B −0.033 −0.350 D2_GZ 0.183 0.850* D2_F 0.360 0.326 DS_forward −0.110 0.390 DS_backward −0.231 0.359 DS_all_WP −0.203 0.407 VPT −0.243 0.084 VPT_1_7 0.126 0.100 VPT_8_14 −0.217 −0.033 VPT_F_8_14 0.217 0.033 LPS_10 0.203 −0.237 LPS_11 −0.707* −0.247 Concerts 0.427 0.536 Practice P 0.467 0.300 Practice MI 0.109 −0.234 Lessons (taking) 0.500 0.417 Lessons (giving) −0.201 −0.092 SR P 0.366 0.522 SR MI 0.305 0.009 Performance_AS −0.103 −0.462 Performance_AP −0.388 −0.650T Note. Significant correlations ($p < .05$) are marked with *, trends ($p < .1$) are marked with T (before Bonferroni correction). Abbreviations: RT_Visual: visual reaction times; RT_Auditory: auditory reaction times; TMT_A: trail making test part A; TMT_B: trail making test part B; D2_GZ: all marked signs of D2 test; D2_F: all false marked signs of D2 test; DS_forward: number of forward digits recalled in digit-span task; DS_backward: number of backward digits recalled in digit-span task; DS_all_WP: age-corrected value of all recalled digits; VPT: score of the visual pattern test; VPT_1_7: score of the patterns 1-7 of the visual pattern test ; VPT_8_14: score of the patterns 8–14 of the visual pattern test ; VPT_F_8_14: errors of the patterns 8–14 of the visual pattern test; LPS_10: number of correct responses for subtest 10 of LPS; LPS_11: number of correct responses for subtest 11 of LPS; SR: sight reading; P: Piano; MI: Minor Instrument; Performance_AS: grade for the performance at sight; Performance_AP: grade for the performance after rehearsal   time to complete the test: the shorter the time to complete the test, the larger the size of the EHS. Another trend was found between performance and EHS: the performance after rehearsal correlated with the number of beats with a value of –0.650 ($p = .058$). This correlation was negative because a smaller number in grade (e.g. 1 compared to 4) corresponded to a better performance. So, for the measurement after rehearsal, participants with a better performance showed a significantly larger EHS in beats. To test whether participants improved their performance through rehearsal, a one-sided t test compared the rated quality of the measurement at sight with the measurement after rehearsal. A trend was detected, $t(8) = 1.835$, $p = .052$. The mean grade for the first measurement (at sight) was 4.0 ($SD = 1.5$) and the mean grade for the second measurement (after rehearsal) was 3.11 ($SD = 1.76$). Discussion The aim of our study was to investigate different factors that might be related to the duration of the EHS of pianists. These were (a) practice, (b) playing tempo, (c) complexity and structure of the SR task, and (d) SR and cognitive performance. We mimicked a real-life piano-accompaniment situation with a pre-recorded flute voice and utilized a novel approach with a combined MIDI and eye-tracking technology to precisely determine the EHS. Comparison of the EHS at sight with the EHS after rehearsal The hypothesis that the EHS increases after a rehearsal phase of 30 minutes had to be rejected because the analysis did not yield any significant differences between both measurements. Only a trend for the difference in EHS latency was discovered. This suggests that pianists did not benefit from practising the music at their high level of skill. An explanation for the lack of a rehearsal effect could be that the EHS of these players was already well established through years of practice and therefore might not change after a short rehearsal phase. Another explanation might be that the sheet music was too difficult for the participants. Increasing the EHS would have led to a higher memory load which might have been impossible to process for the musicians. The fact that only a slight improvement (a statistical trend) in the performance after the rehearsal phase was found could be explained by the possibility that the time to practise the music was too short to significantly improve the performance. As a small rehearsal effect was observed for the performance; this might account for the lack of effect for the EHS. Burman and Booth (2009) found an increase for the perceptual span but also found improved performance for the practised passages. A separate analysis of subgroups showed that three types of changes were present after a rehearsal. Three subjects showed a decrease of the EHS, three an increase and three almost no change. The changes – if present at all – were small to moderate. The groups with altered EHS showed a significant difference in the EHS as expressed in beats after the rehearsal. This explains why the group as a whole did not show a significant effect of rehearsal. Considering these complex results for a short rehearsal of 30 minutes, it might be worth exploring the effect of rehearsal on the EHS a bit more detailed in future studies. Comparison of the EHS for different tempi Playing tempo had a significant effect on the duration of the EHS. The EHS in latency was reduced with faster playing tempo, but increased with slower playing tempo. On the other hand, the EHS in beats decreased with slower tempo (but did not increase significantly for the   fast tempo). The results confirmed the hypotheses and therefore the impact of playing tempo on the duration of the EHS. This effect was already found in terms of the EHS in latency (Furneaux & Land, 1999). The present study also found a correlation between the EHS in beats and the playing tempo. However, the number of played notes was not included in this analysis. It could be possible that the musicians played fewer notes during the fast tempo and more notes during the slow tempo. We did not investigate the influence of playing tempo on the number of played notes which seems to be meaningful for future studies. Comparison of the EHS for different complexities The hypothesis that structural complexity influences the EHS was partly confirmed. The results for the EHS in latency did not show a significant effect of structural complexity, but the number of beats was significantly higher for easy measures than for difficult measures. Other studies found a correlation between the anticipation of notes and the complexity of the score (Rayner & Pollatsek, 1997; Wurtz et al., 2009), but no correlation between the EHS in latency and the complexity of the music. This is quite surprising because in these previous studies, the participants played two or three different pieces (one defined as more complex as the other) at their own tempo and this playing tempo correlated with the complexity of the music. A complex part was played slower than an easier part. In the present study, the tempo was controlled by accompanying the flute voice, so no temporal differences were possible and yet no correlation between the EHS and the difficulty of the music could be found. On the other hand, a higher number of beats for the easy than for the difficult measures was found. It could be possible that there was a greater number of notes in structurally difficult measures than in easy measures (true for some but not all measures). Thus, the increased number of notes could decrease EHS as the hypothetical visuo-motor buffer could only store a certain number of notes. For difficult measures containing the same number of notes as in easy measures, the decrease in number of beats could be due to the fact that the material may be more difficult and time-consuming to process than for the easy measures and therefore fewer notes could be stored in the buffer. However, one has to consider that only few data points (eight measures for the difficult and seven measures for the easy complexity, computation of every beat) have been integrated into this analysis. Therefore, in future studies, a larger amount of data for each group (difficult vs. easy) should be collected. Relationship between the EHS and performance, expertise and cognitive skills The data revealed correlations between the EHS and a) mental speed skills, b) shape recognition skills and c) performance (only for measurement with original tempo after rehearsal). The results of other studies (Furneaux & Land, 1999; Sloboda, 1974; Truitt et al., 1997), which showed a correlation between EHS and skill level, resulting in a larger EHS of better players in contrast to bad SR players, could not be confirmed. In the current study, a correlation between the EHS in number of beats and performance was found but only after rehearsal. On one hand, the effect might be very small so that the number of participants was too small to reveal this effect for both measurements. On the other hand, the SR task might have been too difficult for the participants, as could be seen in their grades ('acceptable' for the first trial and 'adequate' for the second trial). Other studies (Kopiez & Lee, 2008; Lee, 2003; Meinz & Hambrick, 2010) showed correlations between SR performance and working memory capacity, mental speed and expertise. The benefits of mental speed on SR achievements are especially stressed (Kopiez & Lee, 2006;   Kopiez & Lee, 2008). Additionally a significant relationship between SR skill and pattern recognition skills was found (Waters et al., 1998). We tested whether a relationship between the EHS and cognitive skills exists and found that the results of the D2 Test significantly correlated with the number of beats and there was a trend for the TMT part A. This proves that the skill needed to successfully pass the D2 Test is positively correlated with the size of the EHS, meaning participants with a high score in the D2 Test showed a larger EHS than participants with a lower score. Such a mental speed task has not been used before in terms of finding correlations between cognitive performance and SR or the size of the EHS. The fact that both values correlate positively, demonstrates that the scanning ability and the skill to extract information are related to a high EHS. Important to bear in mind is that pianists generally demonstrate very high performance levels on tasks similar to piano-playing (Jäncke, Shah, & Peters, 2000; Jäncke, Schlaug & Steinmetz, 1997). For highly complex SR material, mental and psychomotor speed have been shown to be the dominant predictors (Kopiez & Lee, 2006) but it is clear that both strongly correlate if they are both assessed by tests requiring motor performance. The trend for a correlation of TMT part A and EHS could

be explained by the low number of participants. The negative correlation of EHS and shape recognition skills in the LPS subtest 11 is not meaningful because a high EHS value is associated with a low score for the LPS. This means that either this test was not adequate to assess the shape recognition skills of the pianists, or that these skills are not relevant for the size of the EHS. It is probable that the other cognitive tests were not adequate to fully access the coding strategies used to encode the material in the sheet music and therefore no other meaningful correlations could be found. Another explanation could be that the SR material was highly complex; a previous study showed that working memory capacity becomes important for intermediate SR complexities but importance strongly decreases for high SR complexity levels (Kopiez & Lee, 2006). Limitations One major problem of the study was loss of data and the correction of the missing data. Some measures were not included: measures where fewer than five participants produced an EHS were excluded from the statistical analysis. For the interpolation of the empty measures, the mean of that measure from all other participants was taken. During analysis of the EHS it was obvious that variations between participants, and also within participants for different measures, were large. Although we acknowledge that the mean of all other participants is possibly not the optimal value for data correction, and might have elicited artefacts (resulting in significant correlations or hiding correlations), it appeared to as a pragmatic way to deal with lost data. To avoid further loss of data, it is necessary to establish a good eye tracking signal. In this study, the pianists had to place their head on a chin rest, so they had restricted movement, and they were sitting in an unusual playing position. In future studies, it would be good to include a practice session one day before the measurement. This way, participants could get accustomed to the apparatus. For the computation of the EHS, the entity in beats seems to be the most meaningful measure. The number of beats is stable throughout the whole piece of music, whereas the number of notes changes every measure. Especially when it comes to correlations with complexity of the music, the number of notes is not an adequate measure because in most cases, more notes are in a difficult part than in an easier part of the music. Besides that, a computation of the EHS in ms is only reasonable if a metronome is used for the paradigm. As one can see, the EHS in ms correlates with the playing tempo and therefore one should be very careful in computing the EHS for a piece in which deliberately chosen individual playing tempo is   allowed. One should keep in mind that the EHS in beats changes as well for the playing tempo (here significantly for the slow tempo). Therefore it seems useful that clear distinctions between EHS values for slow and fast pieces of music are made. Conclusion In all instances the eyes were clearly ahead of the hands by a varying amount of time. This delay may be needed to recognize, process and store the information in a hypothetical buffer until the motor output is planned and performed. Some of these findings are inconsistent with previous work on eye-movements in musicians, but we demonstrated little evidence for correlations between the duration of the EHS and the performance of the participant and the participant's mental speed ability. A short rehearsal phase does not influence the EHS, whereas the playing tempo and the difficulty of the music do. We conclude that the EHS is characteristic for each musician and developed over years of practice. It is independent of a short rehearsal phase and not correlated with overall hours of practice and experience of the musician. As studies concerning the computation of the EHS in musicians are rare, these results can be seen as an encouragement for further research into this fascinating topic of a highly complex cognitive and sensory-motor task.
Acknowledgements The authors would like to thank Dennis Trenner for writing the programs for data acquisition and analysis. Ethics statement This study was approved by the local ethics committee of the University of Bremen. Funding This research received no specific grant from any funding agency in the public, commercial, or notfor-profit sectors. The first

author currently receives a PhD grant from the German National Academic Foundation (Studienstiftung des Deutschen Volkes).

References

Burman, D. D., & Booth, J. R. (2009). Music rehearsal increases the perceptual span for notation Music Perception: An Interdisciplinary Journal, 26(4), 303–320. Brickenkamp, R. (1994): D2 Aufmerksamkeits-Belastungs-Test [D2 Concentration Endurance Test]. 8. Auflage. Göttingen, Germany: Hogrefe. Della Sala, S., Gray., C., Baddeley, A., & Wilson, L. (1997). Visual Pattern Test. Oxford, UK: Pearson Assessments. Furneaux, S., & Land, M. F. (1999). The effects of skill on the eye-hand span during musical sight-reading. Proceedings. Biological Sciences / the Royal Society, 266(1436), 2435–2440. Gilman, E., & Underwood, G. (2003). Restricting the field of view to investigate the perceptual spans of pianists. Visual Cognition, 10(2), 201–232. Gruber, H., Jansen, P., Marienhagen, J., & Altenmüller, E. (2010). Adaptations during the acquisition of expertise. Talent Development & Excellence, 2(1), 3–15. Horn, W. (1983). L-P-S Leistungsprüfsystem [L-P-S Performance Testing System]. 2. Auflage. Göttingen, Germany: Hogrefe. Jäncke, L., Schlaug, G., & Steinmetz, H. (1997). Hand skill asymmetry in professional musicians. Brain and Cognition, 34(3), 424–432.   Jäncke, L., Shah, N. J., & Peters, M. (2000). Cortical activations in primary and secondary motor areas for complex bimanual movements in professional pianists. Brain Research. Cognitive Brain Research, 10(1–2), 177–183. Kinsler, V., & Carpenter, R. H. (1995). Saccadic eye movements while reading music. Vision Research, 35(10), 1447–1458. 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. Lee, J. I. (2003). The role of working memory and short-term memory in sight reading. In R. Kopiez, A. Lehmann, I. Wolther, & C. Wolf (eds.), Proceedings of the 5th International Conference of the European Society for the Cognitive Sciences of Music (ESCOM), Hanover, Germany, pp. 121–126. Lehmann, A. C. (2005). Vomblattspiel und Notenlesen – Visuelle Wahrnehmung und Augenbewegungen & Studien zu Gedächtnisprozessen beim Vomblattspiel [Sight-reading and reading music – Visual perception and eye movements & studies about memory processes during sight-reading]. In T. H. Stoffer & R. Oerter (eds.), Allgemeine Musikpsychologie [General Psychology of Music], pp. 883–890. Göttingen, Bern, Toronto & Seattle: Hogrefe. 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: A Journal of the American Psychological Society / APS, 21(7), 914–919. 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. Rayner, K., & Pollatsek, A. (1997). Eye movements, the eye-hand span, and the perceptual span during sight-reading of music. Current Directions in Psychological Science, 6(2), 49–53. Reitan, R. M., & Wolfson, D. (1985). The Halstead-Reitan neuropsychological test battery. Tucson, AZ: Neuropsychology Press. Sloboda, J. (1974). The eye-hand span – an approach to the study of sight reading. Psychology of Music, 2(2), 4–10. Tewes, U. (1994). HAWIE-R. Hamburg-Wechsler-Intelligenztest für Erwachsene, Revision 1991; Handbuch und Testanweisung [Hamburg-Wechsler-Intelligence-Test for Adults, Revision 1991; Manual and Test Instruction]. Bern: Verlag Hans Huber. 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., Townsend, E., & Underwood, G. (1998). Expertise in musical sight reading: A study of pianists. British Journal of Psychology, 89(1), 123–149. 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. Wolf, T. (1976). A cognitive model of musical sight-reading. Journal of Psycholinguistic Research, 5(2), 143–171. Wurtz, P., Mueri, R. M., & Wiesendanger, M. (2009). Sight-reading of violinists: Eye movements anticipate the musical flow. Experimental Brain Research, 194(3), 445–450.

The purpose of this study was to compare the effect of cognitive chunking techniques among first-semester group-piano music majors. The ability to group discrete pieces of information into larger, more meaningful chunks is essential for efficient cognitive processing. Since reading keyboard music and playing the piano is a cognitively complex task, the ability to recognize patterns and execute the necessary motor skills is critical for proficiency in sight-playing. Three groups of beginning group-piano students ($N = 43$) worked on 12 sight-reading examples during six class sessions. A control group rehearsed the examples without benefit of rhythm or pitch drills. Two experimental groups drilled either rhythm or pitch patterns prior to practicing the sight-reading exercises. Following the treatment phase, all subjects were tested on three sightreading examples to determine whether rhythmic and pitch chunking drills impacted reading performance at the keyboard. Subjects were evaluated on rhythmic accuracy, pitch accuracy and continuity. There were significant improvements from pretest to posttest in several subcategories. The pitch experimental group improved significantly in pitch, rhythm and continuity accuracy. The rhythm experimental group improved in rhythm and continuity, while the control group improved only in pitch accuracy. Experimental subjects reported significant engagement in rhythm and pitch chunking even if lack of motor skills impaired the actual performance.