Towards a general model of skills involved in sight reading music

Article *in* Music Education Research · March 2008

DOI: 10.1080/14613800701871363

Reinhard Kopiez
Hochschule für Musik, Theater und Medien Hannover

Music Education Research
Vol. 10, No. 1, March 2008, 41–62

Routledge
Taylor & Francis Group

Towards a general model of skills involved in sight reading music

Reinhard Kopiez* and Ji In Lee

Institute for Research in Music Education, Hanover University of Music and Drama, Hanover, Germany

Sight reading is a functional skill which is essential for all musicians involved in particular fields of western classical music culture. In the last decade, expertise theory has shown that time spent on activities is a good predictor for later achievement in a domain. However, this study is based on the hypothesis that general and elementary cognitive skills as well as practice-related skills must be considered in the explanation of outstanding sight-reading achievement. Different component skills involved in sight reading were attributed to one of three categories: (a) general cognitive skills (working memory, short-term music memory, short-term numerical memory, and Raven's D Matrices); (b) elementary cognitive skills (speed tapping, simple reaction time, trilling speed, and speed of information processing); (c) practice-related skills (practicing solo, sight reading, and inner hearing skills). This resulted in a total set of 23 predictors. These tests were used to evaluate the potential correlates of sight-reading ability. A pre-recorded pacing melody paradigm was used for the sight-reading tasks. Multiple regression analysis revealed that the best combination of predictors is trilling speed, sight-reading expertise acquired up to the age of 15, speed of information processing, and inner hearing. These four predictors can explain 59.6% of variance. Excellence in sight reading is, therefore, the result of a combination of components assumed to be practice-related (sight-reading expertise and inner hearing) and practice-unrelated (speed of information processing). Trilling speed is interpreted in terms of an intersection between task-specific training and practice-independent advantages in movement speed. Our proposed 'general model' of sight reading is the completion of our 'dynamic model' that appeared in an earlier issue of this journal.

Keywords: sight reading; music performance; inner hearing

Introduction

The unrehearsed performance of music, so-called sight reading, is a skill required in particular fields of western music culture. It is characterised by great demands on the performer's capacity to process highly complex visual input (the score) under the constraints of real time – without the opportunity of error correction. It is not only of particular interest for musical professionals, such as piano accompanists, conductors, or orchestra players, but it is also one of the five basic performance skills every musician should acquire (McPherson 1995; McPherson and Gabrielsson 2002). McPherson (1994), and McPherson, Bailey, and Sinclair (1997) defined these skills as follows: the ability to perform a repertoire of rehearsed music, to perform

*Reinhard Kopiez. Email: kopiez@hmt-hannover.de

ISSN 1461-3808 print/ISSN 1469-9893 online

DOI: 10.1080/14613800701871363
http://www.informaworld.com

music from memory (where music was memorised using notation and then re-created aurally), to play by ear (where music was both learned and reproduced aurally), to improvise in both 'stylistically conceived' and 'freely conceived' idioms, and to sight read music without prior rehearsal. The definition of sight reading is sometimes unclear, especially when the performance of a piece, having already been read a couple of times, is still accepted as *sight reading*. For this study, we began with Wolf's definition of sight reading as being a skill of playing music from a printed score or part for the first time without the benefit of practice (Wolf 1976).

This basic assumption that sight reading is a complex transcription task involving a series of overlapping perceptual, cognitive, and motoric processes, which can best be described by a *component skill model,* was first introduced by Waters, Townsend, and Underwood (1998). *Component skills* refer to a selection of independent predictors revealed by multiple regression analysis. The use of component models is a widely accepted approach in cognitive psychology, for example, in determining the processes involved in working memory (Baddeley and Logie 1999). Waters, Townsend, and Underwood (1998) used a set of six predictors to show that sight-reading achievement can be explained by three component skills: pattern recognition in musical score elements, prediction skills, and the ability to use auditory representation (i.e. inner hearing).

Since the early studies by Bean (1938) and Jacobsen (1941), there have been numerous attempts to search for variables that can explain differences between accomplished and less-accomplished music readers. Based on the findings, one can surmise that good readers are able to grasp patterns of notes (instead of single notes), tend to have a larger eye–hand span (thus reading farther ahead), tend to correct notational misprints according to musical plausibility (the so-called 'proofreader's error'), respond to structural aspects of the notational features, show shorter gaze fixations and a higher number of regressive saccades, use auditory imagery, and show a higher amount of accumulated practice in their domain (for a review see Lehmann 2005; McPherson and Gabrielsson 2002; Sloboda 1984). From these studies we can conclude that multiple factors influence sight-reading achievement and can also be used for classification of subjects into performance classes (see Kopiez et al. 2006). Some of the predictors can be clustered and divided into a first group related to features of information processing (e.g. pattern perception), a second group determined by features of eye-movement behaviour, and a third group related to practice (expertise, inner hearing).

The present investigation extends previous research by considering a large number of potential influences on sight reading, ranging from elementary skills (i.e. reaction time) over general cognitive skills (i.e. features of general mental capacity) up to higher cognitive skills (i.e. memory capacity). For the group of general cognitive skills, our selection of music-specific and non-music-specific memory skills was derived from the typical demands of the task. From Sloboda's (1974) early studies on the importance of eye–hand span we know that the ability to read ahead while playing unrehearsed music is a condition for successful sight reading. However, we can assume that without a sufficient short-term memory buffer, the advantage of an extended eye–hand span remains useless. Surprisingly, there are only a few studies that have investigated the positive relationship between short-term memory and sight-reading achievement (Eaton 1978; Waters, Townsend, and Underwood 1998). Waters, Townsend, and Underwood (1998) measured the recall accuracy of briefly presented chords and found a correlation of $r = .64$ with sight-reading achievement. As Waters, Townsend, and Underwood (1998) concluded, sight reading is primarily determined by an efficient input process during the translation of the score into performance movements. Although sight reading makes high demands on working memory, previous research has not examined the relationship between working memory and sight-reading ability. Therefore, we included music-specific as well as non-music-specific tests for short-term and working memory in our study. Based on the findings by Salis (1978), we also included aspects of general mental capacity by using a subsection of the Raven matrices (Raven 2000). Salis found a correlation of $r = .57$ for the total IQ score and sight-reading achievement, and a correlation of $r = .39$ for musical short-term memory and sight reading.

For the group of elementary cognitive skills, selection of predictors was based on the assumption that under the time constraints of sight reading, the speed of information intake and processing plays a crucial role. For example, as Eaton (1978) could show, psychomotor skills (measured by the speed of key identification) are important predictors of sight-reading achievement. Thompson (1985) also regarded sight reading as a transcription task, and he assumed that speed of information processing and reaction time play a crucial role. He included a musical reaction time task and found a correlation of $r = -.54$ between the number of correctly performed measures per second of the Watkins–Farnum Performance Scale and the time needed to perform a note suddenly appearing. His findings are supported by Waters, Townsend, and Underwood (1998): The authors found a correlation of $r = -.42$ between speed needed for a pattern matching time and sight-reading skills and of $r = -.52$ between note-naming time and sight-reading achievement. Perceptual speed also seems to play an important role in sight reading. We therefore considered simple reaction time and psychomotor movement speed (wrist tapping and speed trilling) as well as speed of information processing as measured with the number combination test (NCT) (Oswald and Roth 1997).

For the group of expertise-related skills, we had to consider the acquired sight-reading expertise. Following Lehmann and Ericsson (1993, 1996), we used a pacing voice paradigm (the pianist had to accompany a pre-recorded flute melody which was played metronomically) to measure sight-reading skills, and we conducted retrospective interviews to interpolate the amount of accumulated sight-reading practice. Auditory imagery was considered because the study by Schleuter (1993) revealed some correlation between audiation and sight-reading achievement ($r = .25$). Waters, Townsend, and Underwood (1998) found a correlation of $r = .52$ between the results from an auditory imagination task (visual–auditory matching task) and sight-reading skills. Auditory imagery (also called audiation or inner hearing), developed by Gordon (1986, 1990, 1993), is the ability to imagine the sound of musical notation independent of external sound sources, such as an instrument.

This study is also a completion of our previously proposed *dynamic model* of sight reading (Kopiez and Lee 2006). In this dynamic model we could show that the significance of selected predictor variables changed with task demands on five levels of sight reading: When sight reading was easy, general pianistic expertise was sufficient to be able to excel. With increasing task difficulty, psychomotor speed, speed of information processing, inner hearing, and sight-reading expertise became more important. When sight-reading complexity reached its highest level, sight-reading expertise still remained important, but psychomotor speed (as indicated by trilling speed) became the dominant predictor.

To summarise, the causes for differences in sight-reading achievement are diverse and cannot be explained exclusively by differences in basic features of music reading (e.g., eye movements, pattern perception, etc.). We propose that a general model of sight reading must consider relevant factors from all levels of information processing, including variables influenced by training as well as those independent of practice. Our study is thus based on the main hypothesis that sight-reading achievement is the result of a combination of variables related to general cognitive skills, elementary cognitive skills, *and* expertise. In other words, we propose that general and elementary cognitive skills as well as practice-related skills must be considered in the explanation of outstanding sight-reading achievement.

Rationale for the study

This study had two aims: first, to develop a general model of sight reading which can explain observed achievement variance with a limited number of predictor variables covering many levels of information processing; second, to give insight into the underlying predictor structure by identifying basic components and factors that group selected predictors.

Method

Subjects

We decided to use the piano as our test instrument due to the availability of MIDI-based methods of recording and objective procedures of performance evaluation. Thus, 52 piano major students or graduates from the Hanover University of Music and Drama served as subjects for this experiment. These included undergraduates, postgraduates, professional accompanists and winners of sight-reading competitions. The mean age was 24.56 years (SD = .49); 24 males and 28 females took part.

Sight-reading material and procedure

Sight-reading stimuli were selected from the University of South Africa exam syllabus for piano sight reading (UNISA 1995). The advantage of using this material was that the pieces were already assessed and categorised into increasing levels of complexity. We decided on five pieces with increasing complexity that were of similar length and were arranged for solo voice and bi-manual piano accompaniment by a professional composer. External judges (professional piano accompanists) evaluated the different levels of task complexity. Additionally, the complexity of pieces was scrutinised for physical surface complexity (see Table 1). The pre-recorded pacing melody paradigm (Lehmann and Ericsson 1996) was used to create time constraints by keeping the tempo constant. Subjects had to accompany a pre-recorded solo voice performed by a violinist who had recorded these melodies while synchronising with a metronome through headphones (for an example see Appendix 1). The solo voice was presented through a loudspeaker at a comfortable sound pressure level. Subjects were given 60 seconds to look at each piece and tempo. The cue to start playing was indicated by two full bars of clicks before the piece. A MIDI keyboard (with

Table 1. Physical surface complexity of the sight reading stimuli of each task level.

Level	Left hand	Right hand	Both hands	No. of bars	Average no. of notes in one bar	Total duration (s)	Average time for one bar (s)
1	80	89	169	23	7.34	45	1.95
2	98	90	188	20	9.4	48	2.45
3	103	188	291	37	7.86	52	1.45
4	93	105	198	21	9.42	50	2.38
5	188	222	410	21	19.52	92	4.38

Note: Rank order of task levels is based on expert ratings.

weighted hammers) was used to record the performances of subjects directly into the sequencing programme CUBASE 5.1 (Steinberg 2000). Subjects were first given two warm-up pieces to become familiar with the laboratory situation and the test procedures. The entire procedure lasted about three hours.

Predictor variables: measurement of selected skills

The 23 independent variables selected were divided into three categories: general cognitive skills, elementary cognitive skills, and practice-related skills (Table 2).

General cognitive skills

Working memory. Subjects were required to add or subtract in steps of one to an increasing number of digits in a 3×3 matrix. The displayed matrix started off with only two cells being active, but ended with seven active cells. Subtraction or addition was indicated by four up or down arrows appearing in a random order. An arrow pointing upwards meant plus one and an arrow pointing downwards meant minus one. Each calculation had to be performed on the current value of the cell, and the result had to be remembered. This task consisted of five warm-up and 18 test exercises. For each sequence the percentage of correct answers (all or none criterion) was calculated with software developed by Oberauer et al. (2000) and used for data analysis.

Short-term memory (STM) test. This numerical test consisted of a series of numbers, shown one at a time on a computer screen, which the subjects had to remember. Answers were typed in by the investigator and were shown on the screen with the option of being corrected by the subjects. Feedback was given as to how many of the numbers were remembered correctly. The test started with four digits per task and then increased by one digit every time until it reached nine digits per task. There were two warm-up and 18 test exercises. For scoring, the percentage of correct answers for each sequence was used (all or none criterion). Measurement was performed with software developed by Oberauer et al. (2000).

Short-term music-specific memory (STMM) test. This task was an alternative to the numerical STM test and tested music-dependent short-term memory. Tasks were derived from a study by Drösler (1989). Subjects had to look at a short melody of

Table 2. List of 23 grouped predictor variables and methods used for measurement.

Skills group	Predictor variable (name used for data analysis)	Method of measurement
1. General cognitive skills	Short-term memory capacity (STM)	Researcher-developed software (mean% of correct items)
	Working memory capacity (WM)	Researcher-developed software (mean% of correct items)
	Short-term music-specific memory (STMM)	Researcher-developed software (no. of performed notes)
	General mental capacity (Raven D)	Series D of Raven's SPM (no. of correct items)
2. Elementary cognitive skills	Speed of information processing (NCT)	Number combination test (duration in seconds)
	Simple visual reaction time (RTV)	Researcher-developed software (median in ms)
	Simple auditory reaction time (RTA)	Researcher-developed software (median in ms)
	Tapping speed	Tapping device (median in Hz of both hands)
	Trilling speed* over 15 s, f.c.** 1-3, average of two trials (Trill 1-3)	Keyboard trill (median in Hz)
	Trilling speed* over 15 s, f.c.** 3-4, average of two trials (Trill 3-4)	
3. Expertise-related skills	Accumulated hours of solo practice up to the age of 10, 15, 18 and total (Solo 10, 15, 18, total)	Retrospective interview
	Accumulated hours of piano lessons up to the age of 10, 15, 18, and total (Lessons 10, 15, 18, total)	Retrospective interview
	Accumulated hours of sight-reading expertise up to the age of 10, 15, 18, and total (SR. 10, 15, 18, total)	Retrospective interview
	Inner hearing	Embedded melodies test (d')

Note: *All trills were played with the right hand; **f.c. = finger combination.

12 bars for 1 minute and then try to play as many correct notes as they could from memory on a MIDI keyboard. The performance was recorded. Although it is known that encoding of unknown melodies is a multidimensional process which considers all musical parameters available, pitch and contour information are a better memory cue than rhythm (Hébert and Peretz 1997; Sloboda and Parker 1985; White 1960). Therefore, only the sum of all correct pitches performed by the subjects (omitted notes were allowed) was used for data analysis.

Raven's D Matrix. To check for influences of general mental capacity, we used the Standard Progressive Matrices (Raven 2000). Due to the limited time, only series D (12 items) was used. Scoring was based on the number of correct items.

Elementary cognitive skills

Number combination test (NCT). In this pencil-and-paper test of perceptual and processing speed, subjects had to join the numbers 1–90 in ascending order as quickly as possible. The time taken to complete this test was measured. Test scores indicate speed of information processing (Oswald and Roth 1997).

Auditory and visual reaction time (RTA and RTV). Simple reaction time was measured using auditory and visual cues. Subjects had to release a morse key as soon as they saw or heard the stimuli. The time interval of stimulus appearance varied randomly between 500 and 2000 ms after a key was pressed by the subject. Data was recorded using researcher-developed software and the median for each modality in ms was used for the data analysis. There were five warm-up and 20 test exercises for each modality.

Speed trilling. The music-specific psychomotor movement task consisted of speed trilling for 15 seconds. The task used the following two types of speed trills: the thumb and middle finger of the right hand on C4 and D4 (Trill 1-3); the middle finger and ring finger on E4 and E5 (Trill 3-4). Both trills were repeated once, and the average of both medians in Hz was used for the data analysis.

Speed tapping. The non-music-specific psychomotor speed task was speed tapping (wrist tapping) for 30 seconds on a morse key. Subjects were given a test trial, and the start hand was allocated randomly. Researcher-developed software was used for evaluation, and the median of the inter-tap interval (ITI) for both hands in Hz was calculated.

Practice-dependent skills

Inner hearing. We used the *embedded melody paradigm* (the forced choice method developed by Brodsky et al. 2003) to test for auditory imagery. Pre-existing variations from piano literature were used by combining the original theme with a variation written by the composer or with a so-called 'lure melody' written by a composer of our department (see Appendix 2). The lure melody was similar to the theme but contained distinct differences and had a significant deviation from the underlying melodic or harmonic structure. Thus, for each example there were three versions: theme, variation and the lure melody. First, the variation of each theme was shown for 45 seconds using a PowerPoint presentation. Second, subjects had to imagine the sound without humming or singing the score. Third, the presentation was followed: (a) by the theme; or (b) by the lure melody through the speakers. Finally, subjects had to decide whether or not the theme was embedded in the variation seen (forced choice paradigm). Sound examples could be repeated. The *d* prime value was then calculated (MacMillan and Creelman 1991). After a pre-test, it was decided to use two warm-ups and five samples.

Retrospective interviews. We conducted retrospective interviews (see Lehmann and Ericsson 1996) about the number of accumulated hours of practice for solo and sight reading per week and the number of years of piano lessons starting from the beginning of instrumental lessons. Based on these interviews, we calculated the number of accumulated hours of practice for chamber music, accompaniment, sight reading and solo up to the age of 10, 15, 18 and total. Sight-reading expertise (SR) was defined as the sum of reported practice time for chamber music, accompaniment, and sight reading over 52 weeks per year.

Results

Scoring

Scoring of sight-reading achievement

Since the scoring of the recorded performances was critical, we considered numerous possible methods of evaluation. For example, Kornicke (1995) used expert ratings on a 32-item sight-reading performance scale for the quality of dynamics, pedalling, interpretation, and the number of pitch and rhythm errors. We decided to use an evaluation method that is: (a) objective, through the use of transparent and adjustable evaluation criteria; (b) repeatable and independent of the raters' reliability; and (c) based on an automised procedure. The only way to fulfil these three conditions was with a software-based solution that counts the number of correctly performed pitches within a given timeframe as the achievement criterion.

Data analysis of sight-reading achievement was done using the software *Midicompare* developed by Dixon (2002). This programme outputs the score matches (matched notes) for each hand which were played by subjects within the chosen window of time (.25 seconds) in each direction. The influence of window size (time before and after a note event) on the evaluation of performance was determined in a pre-test. Our decision to use a small time window for performance evaluation was based on trials with time windows of different sizes: A time window larger than 250 ms mainly increased the number of unmatched and extra notes. In this study, only the percentage of total score matches for both hands were considered.

Results of sight-reading performances

Table 3 shows the sight-reading achievement results for all five levels. However, this paper deals with the total score only. The changing influence of predictor variables at different task levels was part of a previous study (Kopiez and Lee 2006; Lee 2004). The minimum score was 27%; the maximum score was 97% with a mean score of 61.55% (SD = 17.34%).

Table 3. Table of the scores of the sight-reading achievement from level 1 to level 5 and the total performance scores.

Levels	Min. (%)	Max. (%)	Mean (%)	SD (%)
1	50	100	87.95	14.23
2	28	96	80.38	16.66
3	21	99	72.12	22.96
4	8	99	49.42	27.63
5	8	95	39.50	23.10
Total average	27	97	61.55	17.34

Correlations between the total sight-reading achievement and the 23 predictor variables

Correlation analysis (Table 4) showed that Speed Trill 3-4 had the highest correlation with total sight-reading achievement, followed by SR 15, Speed Trill 1-3, NCT, Inner Hearing, SR 10 and 18 and Solo 10. These eight predictor variables correlated significantly with the total sight-reading achievement. The most intriguing finding was the importance of psychomotor movement speed (as indicated by trilling speed) and speed of information processing (as indicated by the NCT).

Multiple regression analysis

Because of the inter-dependencies of predictors, all 23 predictor variables were used as predictors for the multiple regression analysis. The results of the analysis (Table 5) show four models, of which the first encompassed only the median in Hz of Speed Trill 3-4 for 15 seconds. This predictor alone could explain only 31% of total sight-reading achievement, and the finding is peculiar. Speed trilling could reflect additional skills that are not practice-related as well as indicate a characteristic of the information processing system to optimise motor skills and motor planning under time constraints. The second model included the number of accumulated hours of sight-reading expertise up to the age of 15; these two variables could explain 47% of total sight-reading achievement. The third model consisted of Speed Trill 3-4,

Table 4. Correlations between the total sight-reading achievement and the 23 predictor variables ($n = 52$).

Rank	Predictor variables (variable names)	r	p	r^2 (%)
1	Speed trill 3-4	.560	.000**	31.36
2	SR expertise up to 15 (SR 15)	.496	.000**	24.60
3	Speed trill 1-3	.445	.001**	19.80
4	Number combination test (NCT)	−.440	.001**	19.36
5	Inner hearing (d')	.427	.002**	18.23
6	SR expertise up to 10 (SR 10)	.401	.002**	16.08
7	SR expertise up to 18 (SR 18)	.359	.009**	12.88
8	Solo up to 10 (Solo 10)	.311	.025*	9.67
9	Working memory (WM)	.261	.062	6.80
10	Solo up to 15 (Solo 15)	.252	.071	6.35
11	SR expertise total (SR total)	.251	.073	6.30
12	Solo total (Solo total)	.204	.146	4.16
13	Short-term music memory (STMM)	.191	.175	3.64
14	Solo up to 18 (Solo 18)	.160	.233	2.56
15	Years of piano lessons up to 10 (Lesson 10)	.184	.192	3.38
16	Years of piano lessons up to 15 (Lesson 15)	.183	.194	3.34
17	Short-term memory (STM)	.178	.209	3.16
18	Speed tapping (Tapping)	.122	.388	1.48
19	Reaction time (visual) (RTV)	−.151	.285	2.28
20	Years of piano lessons up to 18 (Lesson 18)	.101	.474	1.02
21	Raven's D Matrices (Raven)	.118	.405	1.39
22	Years of piano lessons total (Lesson total)	.058	.684	.33
23	Reaction time (auditory) (RTA)	.018	.897	.018

Note: *$p < .05$; **$p < .01$ (two-tailed).

Table 5. Regression analysis showing the influence of 23 predictor variables on total sight-reading performance (R^2 adjusted = .59).

	Variables	R^2 adjusted	ΔR^2	Standardised beta coefficient	t	p
1	Trill 3-4	.312		.570	4.90	000
2	Trill 3-4	.472	.160	.513	4.99	.000
	SR 15			.414	4.03	.000
3	Trill 3-4	.560	.080	.485	5.15	.000
	SR 15			.388	4.12	.000
	NCT			−.307	−3.27	.002
4	Trill 3-4	.596	.030	.422	4.46	.000
	SR 15			.380	4.20	.000
	NCT			−.258	−2.79	.008
	Inner hearing			.221	2.29	.026

Note: Method for variable entry: Stepwise. Criterion to enter a new variable was $p < .05$ and for removal was $p > .10$.

SR 15 and number of seconds taken to complete the NCT; these three variables could explain 56% of total sight-reading achievement. Model 4 was the optimum solution of the stepwise multiple regression analysis and consisted of Speed Trill 3-4, SR 15, NCT and Inner Hearing achievement. These four predictor variables resulted in an adjusted R^2 of .59 and could explain up to 60% of total sight-reading achievement. Beta coefficients for NCT indicated a negative relationship between sight-reading achievement and NCT, which means that higher speed of information processing results in higher sight-reading achievement. The quality of the obtained regression model was evaluated using a scatterplot between observed and predicted performance values (see Figure 1). As Figure 1 shows, the model reflected the observed data to a high degree.

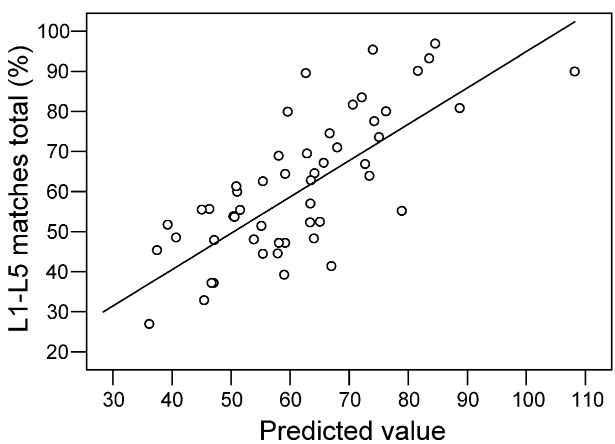

Figure 1. Evaluation of the predictive quality of the general model of sight reading, resulting from multiple regression analysis (R^2 adjusted = .59; x-axis: sight-reading achievement as predicted by the regression equation: y-axis: observed achievement as total percentage of all score matches for five different levels of task difficulty).

The role of deliberate practice in sight-reading achievement

To consider the influence of domain-specific practice, subjects were allocated to groups through a median split of the total achievement score. Ericsson, Krampe, and Tesch-Römer (1993) found an increasing difference between highly skilled violin experts and amateurs in the number of accumulated hours of instrumental practice up to the age of 18. Our results from the retrospective interviews also revealed a significant difference in the accumulated hours of sight-reading expertise between the upper and lower 50% of sight-reading performers. As Table 6 shows, there existed a significant difference from the age of 10 onwards, and this continued into the total accumulated number of hours of sight-reading practice. By age 10, the lower 50% had accumulated only 18 hours compared to 183 hours from the upper 50%. By age 15, the lower 50% had accumulated 199, but the upper 50% had accumulated 710 hours, and by 18, the lower 50% had accumulated 634 compared to 1262 hours from the upper 50%. Differences were significant at all age levels.

Factor analysis of predictor structure

A basic assumption of this study was that sight reading must be explained as a result of many predictors which can be divided into at least three component groups: general cognitive skills, elementary cognitive skills, and expertise-related skills. Our division of the selected 23 predictor variables into these three groups was based on intuition. However, a valid test of the underlying component structure and the relationship between the selected predictors could only be made through principal component analysis (PCA), also known as factor analysis. As an important difference from the regression analysis, the PCA does not look for a model fit between predictors and dependent variables, but for the relationship between predictors. This would give us insight into the underlying component structure of the predictors considered. The method used was PCA with varimax rotation (see Table 7).

The PCA revealed eight factors with eigenvalues larger than 1 that explained 82.26% of variance. Factor I is labelled the ‘piano lessons’ factor and explained about 14% of variance. It consisted of the years of piano lessons up to a particular age, but excluded the total years of lessons. Factor II also explained about 14% of variance and is labelled the ‘sight-reading factor’. It included all variables related to sight-reading expertise. Factor III comprised all expertise in solo playing and

Table 6. The average number of accumulated hours of sight-reading expertise (SR) up to the ages of 10, 15, 18 for the upper and lower 50% of sight-reading performers.

	Upper 50% of SR performers				Lower 50% of SR performers					
	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.	t	p
SR 10	183	305	0	1126	18	43	0	182	−2.59	.008**
SR 15	710	821	0	3520	199	316	0	1248	−2.96	.005**
SR 18	1262	1447	30	6528	634	757	0	2522	−1.96	.05*
SR total	3327	3942	108	13,962	1690	1937	15	7002	−1.90	.05*

Note: Grouping into upper and lower performers was based on median split of the total SR performance (*p <.05; **p <.01, two-tailed).

Table 7. Factor loadings of rotated component matrix of all 23 predictor variables (factor loadings <.40 have been omitted).

	Piano lessons	Sight reading	Solo	Cognitive (speed/memory)	Reaction time	Psycho-motor speed/ Music memory	General piano skills	Innner hearing
Factor	I	II	III	IV	V	VI	VII	VIII
% of variance	14.2%	13.9%	13.8%	9.6%	8.4%	8.1%	8.1%	5.9%
SR 10		.848						
SR 15		.965						
SR 18		.946						
SR total		.719					.532	
Solo 10			.798					
Solo 15			.888					
Solo 18			.906					
Solo total			.666				.520	
Lessons 10	.957							
Lessons 15	.964							
Lessons 18	.937							
Lessons total							.890	
Inner hearing								.882
STMM						.663		
NCT				−.645				
Raven D				.603				
STM				.781				
WM				.834				
RT picture					.898			
RT sound					.872			
Tapping speed						.616		
Trill 1-3 (mean)			.415			.592	.410	
Trill 3-4 (mean)						.730		

Note: Extraction method: Principal component analysis. Rotation method: Varimax with Kaiser normalisation. Total variance explained: 82.26%. For explanation of variable names see Table 2.

explained 13.8% of variance. It included all predictors relevant for the amount of solo practice. Thus, it is labelled the 'solo factor'. Factor IV comprised variables related to general cognitive skills (Raven), speed of information processing, and memory. The negative factor loading of NCT meant that a small value in the mental speed task corresponded to a high value in the Raven matrices and the memory tasks. It explained 9.6% of variance and is labelled the *cognitive factor*. Factor V showed a clear picture and was represented by simple reaction times. It is called the *reaction time factor* and explained 8.4% of variance. Factor VI included tapping speed for both hands as well as trilling speed for both finger combinations. However, music-specific short-term memory (STMM) was also related to this factor. It is labelled the *psychomotor speed factor* and explained 8.1% of variance. Factor VII revealed a heterogeneous picture: Included are the total sight reading and solo experience, total years of piano lessons, and speed trilling with finger combination 1-3, explaining 8.1% of variance. This factor showed that sight-reading expertise was represented separately (Factor II) but was also part of general piano skills. Factor VIII was the last factor found and explained 5.9% of variance. It is called the *inner hearing factor* because the scoring in the inner hearing task was the only variable loading highly on this factor.

Discussion

The results of this study indicate that selected elementary cognitive skills and practice variables are potential correlates of sight-reading ability. Significant relationships were found in correlation and multiple regression analysis. Nevertheless, this study was guided by the hypothesis that sight reading is a complex skill and can only be explained if practice-dependent as well as non-practice-dependent predictors are considered. Thus, in the next paragraphs, we will discuss the contribution of the three groups of predictors (general cognitive skills, elementary cognitive skills, and expertise-related skills) to the explanation of differences in sight-reading achievement.

Evaluation of performance

Our intention in using a purely quantitative score-matching paradigm with the aim of note-perfect sight reading for the evaluation of performance was to obtain an objective and reliable measurement for sight-reading achievement. However, the adjustment of the time window for counting score matches to ± 250 ms was a restrictive condition resulting in an underestimation of score matches. For example, matches were not counted if the performer added expressive deviations to his or her playing and in doing so performed outside the given time frame. Of course, score matches are only one aspect of sight reading, and it is known that good sight readers use intelligent strategies to correct errors or to omit some notes in order to maintain tempo and harmony. These aspects can influence the impression of musical quality of sight reading, but they are difficult to evaluate and would require a different evaluation method (e.g., expert ratings of the musical quality of performances).

General cognitive skills

None of the predictors from this group, such as working memory or short-term memory, was revealed as being significant in multiple regression analysis. Thus our hypothesis of a *bottleneck variable*, comprising memory and general mental capacity, was incorrect. Sight reading does not seem to be influenced by memory capacity or aspects of intelligence. However, we have to bear in mind that the measurement of working and short-term memory capacity was based on number-related tasks. Of course, it would also be possible in future experiments to use verbal or spatial-figural tasks in working memory tests. As Oberauer et al. (2000) could show in their study of the structure of working memory, verbal and numerical working memory are closely related, but spatial-figural working memory is represented by a different factor in factor analysis.

The underlying component structure of predictors as revealed by factor analysis

As the factor analysis revealed, the statistical structure of predictors is best represented by eight factors and not by three groups only (i.e., general cognitive skills, elementary cognitive skills, and practice skills). The component structure derived from factor analysis differs from the initial classification of variables. For example, mental speed (NCT) is more related to general cognitive skills than to elementary cognitive skills and inner hearing builds a separate factor.

The question remains open as to whether or not the three statistical methods used for data analysis (correlation analysis, regression analysis, and PCA) reveal the same information. In the first step, we had to compare the results from statistical methods that search for relationships between independent and dependent variables. Correlation analysis (Table 4) revealed eight significant correlations between predictors and sight-reading achievement. If these significant variables are compared with the results from multiple regression analysis (Table 5), we see that all predictors from the regression analysis can be found on the first five ranks of Table 4. However, correlation analysis does not consider inter-dependencies between independent variables. Thus, the number of significant variables is higher compared with regression analysis. In the second step, we compared these findings with the results from PCA, while keeping in mind that a PCA has a different statistical perspective and looks for the underlying structure of predictor variables and not for the relationship between independent and dependent variables. However, to test for the validity of factors found, correlations between rotated factor scores and the dependent variable (sight-reading achievement) were calculated. A significant correlation between sight-reading achievement and factor scores was found for the following factors: the sight-reading expertise factor (F II; $r(52) = .36$; $p = .00$), the psychomotor speed/trilling factor (F VI; $r(52) = .40$; $p = .00$), and the inner hearing factor (F VIII; $r(52) = .42$; $p = .00$). The only significant independent variable that did not show a correlation between its related factor and sight-reading achievement was speed of information processing (F IV). Thus, we conclude that despite different perspectives and different statistical approaches, all three methods of data analysis used, nearly 'tell the same story' and confirm our findings.

To summarise, the basic assumption of sight reading as a skill which can best be described in terms of a component structure could be confirmed. Waters, Townseend, and Underwood (1998) found the ability for pattern recognition, prediction of subsequent chords, and the ability for auditory imagery (which corresponds with our inner hearing test) to be the best three predictors for sight-reading achievement. In their study, however, components were only found by ANOVA, and not by PCA. This might explain the difference between observed component structure and the results of our study.

Elementary cognitive skills

The most surprising result from multiple regression analysis is the importance of trilling speed and speed of information processing (NCT) as predictors for sight-reading achievement (see Table 5). As Model 4 shows, trilling speed with finger combination 3-4 is a strong predictor. How can this be explained? Sight reading is a highly specialised skill which is based on neuro-cognitive (e.g., mental speed) as well as on practice-dependent prerequisites. From the perspective of training, statistical analysis revealed a correlation of $r = .54$ ($p = .00$, one-tailed) between trill 1-3 and of $r = .24$ ($p = .03$) between trill 3-4 and the total acquired hours of solo practice. From the neuro-cognitive perspective, speed trilling depends on the general tendency to purposive optimisation (e.g., time-critical movements or motor planning). This explanation is supported by a correlation of $r = .22$ ($p = .05$) between speed trill 3-4 and finger tapping speed. To summarise, trilling speed represents the intersection between task-specific training and practice-independent advantages in movement speed.

The predictor speed of information processing (so-called 'mental speed' as measured by the NCT) was revealed as being a significant predictor from the group of elementary cognitive skills. This predictor is not domain-specific for sight reading, and its outstanding importance is a surprise. No previous literature has reported an influence of speed of information processing on music performance. In sum, speed of information processing is the only predictor in the regression model assumed to be independent of practice.

Expertise and practice-related skills

We found that there is a crucial time window for the acquisition of sight-reading expertise. The number of accumulated hours of sight-reading practice up to the age of 15 is the best predictor for sight-reading achievement. This raises the issue of sight reading after the age of 15, as the correlation between sight-reading achievement and sight-reading expertise up to the age of 18 became less significant (see Table 4). Based on the weak correlation between the total sight-reading expertise and achievement ($r = .25$; $p = .072$), we can conclude that the total sum of acquired expertise is not the best or most important predictor. In other words, the questions of *how much* and *when* sight-reading expertise is acquired have to be considered. We would like to emphasise this new developmental perspective in the acquisition of expertise. This finding conforms to the idea of a critical time window for skill acquisition, which is also apparent in other domains (e.g., language). However, as Table 6 shows, there was a considerable standard deviation in the amount of interpolated hours of sight-reading practice. This means that the calculated values of practice time up to a specific age should not be interpreted as absolute practice times, but as an indicator for the distribution of time resources to selected activities (see Lehmann and Ericsson 1993, 1996).

The predictor inner hearing was allocated to the group of practice-related skills. According to Waters, Townsend, and Underwood (1998), this influence of inner hearing is interpreted as a priming effect. In their study, the authors could show that 'auditory imagery' (which can be compared to the ability of inner hearing) was the third best predictor in a simple sight-reading task. We assume that the ability to imagine sound from notation can help generate predictions about what is coming next in the score. This would mean that sight reading is not only a question of eye–hand co-ordination but also of eye–ear–hand co-ordination. In this case, inner hearing would function as an additional information channel that would be useful for anticipatory information processing (Lee 2006). However, we have to bear in mind that our study is based on results from pianists. This means that a transfer of our findings to other instrumentalists could result in a different weighting of variables. For example, due to higher demands in intonation sensitivity, for violinists or singers the variable inner hearing could be much more important than for pianists.

General discussion

This study was guided by the hypothesis that sight reading is a complex skill which can only be explained if one considers variables related to both cognitive, practice-independent skills (e.g., working memory, short-term memory and elementary cognitive tasks) *and* to practice skills. First, the important role of expertise could be confirmed; in other words, sight-reading expertise is determined by the time spent on activities related to this skill. The concept of deliberate practice seems to be a general explanation for skill acquisition (Lehmann 1997). However, we added a new aspect by pointing out that there is a critical time window (before turning 15) for optimal training. Thus, skill acquisition has to start early in order for the person to become an excellent sight reader.

Second, we added speed of information processing as a new predictor. Fluency of sight reading is more influenced by mental speed than by memory capacity or general intelligence. This means that sight reading benefits most of all from mental skills which are time-sensitive, a factor not measured by most intelligence tests. Sight reading poses different demands on processing speed than do other kinds of mental activities in that the information flows through the system so quickly as to avoid 'short-term memory bottleneck'. In sight reading, a lack of mental speed cannot be compensated for by working memory or short-term memory capacity.

The question remains open as to whether mental speed remains constant or changes with experience and practice. In the current state of research there is a clear answer: Mental speed is not affected by training, but it is a preexisting ability. This argument comes mainly from studies on the effects of aging on mental speed. For example, Kail and Salthouse (1994) and Salthouse (2000) argued that speed of processing is a fundamental part of the architecture of the cognitive system as it develops across the entire life span. Another argument comes from intelligence research with monozygotic twins: As Neubauer et al. (2000) argued, genetic analyses suggest that most of the phenotypic correlation between mental speed and intelligence is due to genetic factors. From this we can conclude that speed of information processing facilitates sight reading but that the activities which give rise to better sight reading do not significantly increase performance on the NCT task. We would argue the other way around: Mental speed functions as a selective variable, and sight readers with high mental speed increase their expertise through increased practice. However, correlation between the total amount of accumulated sight-reading expertise and NCT is weak ($r = -.29$, $p = .03$).

Third, psychomotor speed, which is represented by trilling speed, was found to be a strong predictor of sight-reading achievement. As the PCA revealed, trilling speed is represented by a separate factor that also includes tapping speed. An important question remains open: Is trilling speed only a result of piano practice (as a domain-specific speed), or does it contain components of general psychomotor speed? This question could only be answered by a repeated trilling or tapping task and measurement of improvement over a large number of trials. However, up until now, there has been only one study on the improvement of tapping speed through practice. As Peters (1976) could show, short tapping trials of 10 s each over a period of four weeks with more than 1200 tappings improved speed for both hands. Improvement was significantly higher for the non-preferred left hand. After four weeks, the non-preferred hand reached and surpassed the tapping speed of the preferred hand. The question of an improvement in trilling speed through practice cannot be answered by the current state of research. To summarise, differences in psychomotor speed can be explained by training effects as well as by physiological effects, such as a high personal tremor speed (Freund 1989), which is determined by genetic make up. However, this is a question for future research.

In summary, our general model of sight reading predicts that sight-reading achievement can be explained by a linear combination of psycho-motor speed, early acquired expertise, mental speed, and the ability for auditory imagery. General cognitive skills (e.g., memory capacity) do not play a significant role. At a very high level, sight reading is dependent not only on expertise and practice-related variables but also on genetically determined variables, such as mental speed (Kopiez et al. 2006).

Educational implications

In our study we could show that sight-reading achievement is determined by a combination of skills assumed to be practice-related (e.g., sight-reading expertise and inner hearing) and practice-unrelated (e.g., speed of information processing). Of course, speed of information processing cannot be increased by training but this does not mean that it cannot be changed by different approaches. The easiest way to enhance the information processing capacity would be to practice pattern recognition and chunking of note events (for examples of different practice methods of chunking see Lehmann, Sloboda, and Woody 2007).

Acknowledgements

We are indebted to Niels Galley (University of Cologne, Germany) for the reaction time and speed tapping software, to Nicholas Sander and Klaus Oberauer (University of Mannheim, and University of Potsdam, Germany) for the memory capacity software, and to Wolfgang Auhagen for helpful comments on a previous version of this paper. Sound examples, a video of

the procedure, the software MidiCompare and test materials are available from the website http://musicweb.hmt-hannover.de/sightreading.

Notes on contributors

Reinhard Kopiez holds a degree in classical guitar, and Master's and PhD degrees in musicology. From 2000 to 2003 he was vice-president of ESCOM and organiser of the 2003 5th Triennial ESCOM Conference in Hanover, Germany. From 2001 to 2005, he was president of the German Society for Music Psychology (DGM). He holds a position as Professor of Music Psychology at the Hanover University of Music and Drama. His main specialist area is performance research with a special interest in intonation, long-term performance (e.g. Satie's 'Vexation') and sight reading.

Ji In Lee, started playing the piano at the age of six and completed her BMus, BMus (Hon) and MMus in South Africa in piano and in pedagogy. She has completed her PhD in music psychology at the Hanover University of Music and Drama with a study on the sight reading of music.

References

Baddeley, A.D., and R.H. Logie. 1999. Working memory: The multiple-component model. In *Models of working memory. Mechanisms of active maintenance and executive control,* ed. A. Miyake and P. Shah, 28–61. Cambridge: Cambridge University Press.

Bean, K.L. 1938. An experimental approach to the reading of music. *Psychological Monographs* 50, no. 6: 1–80.

Brodsky, W., A. Henik, B. Rubinstein, and M. Zorman. 2003. Auditory imagery from musical notation in expert musicians. *Perception and Psychophysics* 65, no. 4: 602–12.

Dixon, S. 2002. *Midicompare.* Computer software. Vienna, Austria: Austrian Institute for Artificial Intelligence.

Drösler, A. 1989. Visuelle Wahrnehmungen von Notenfolgen: eine experimentelle Untersuchung zum Kurzzeitgedächtnis von Experten- und Laienmusikern verschiedener Altersgruppen [Visual perception of note sequences: An experimental investigation of short-term memory in musical experts and non-experts of different age]. PhD diss., Technische Universität Berlin.

Eaton, J.L. 1978. A correlation study of keyboard sight-reading facility with previous training, note-reading, psychomotor, and memorization skills. PhD diss., Indiana University. *Dissertation Abstracts International-A* 39, no. 7: 4109.

Ericsson, K.A., R.T. Krampe, and C. Tesch-Römer. 1993. The role of deliberate practice in the acquisition of expert performance. *Psychological Review* 100, no. 3: 363–406.

Freund, H.-J. 1989. Handmotorik und musikalisches Lernen [Hand motor performance and musical learning]. In *Musik-Gehirn-Spiel. Beiträge zum vierten Herbert von Karajan-Symposium, Wien 24.–25. Mai 1988,* ed. H. Petsche, 101–10. Basel, Switzerland: Birkhäuser.

Gordon, E.E. 1986. *The nature, description, measurement, and evaluation of music aptitudes.* Chicago, IL: GIA Publications.

———. 1990. *A music learning theory for newborn and young children.* Chicago, IL: GIA Publications.

———. 1993. *Learning sequences in music. Skill, content, and patterns.* Chicago, IL: GIA Publications.

Hébert, S., and I. Peretz. 1997. Recognition of music in long-term memory: Are melodic and temporal patterns equal partners? *Memory and Cognition* 25, no. 4: 518–33.

Jacobsen, O.I. 1941. An analytical study of eye-movements in reading vocal and instrumental music. *Journal of Musicology* 3: 1–32, 69–100, 133–64.

Kail, R., and T.A. Salthouse. 1994. Processing speed as a mental capacity. *Acta Psychologica* 86, no. 2–3: 199–225.

Kopiez, R., and J.I. Lee. 2006. Towards a dynamic model of skills involved in sight reading music. *Music Education Research* 8, no. 1: 97–120.

Kopiez, R., C. Weihs, U. Ligges, and J.I. Lee. 2006. Classification of high and low achievers in a music sight reading task. *Psychology of Music* 34, no. 1: 5–26.

Kornicke, L.E. 1995. An exploratory study of individual difference variables in piano sight-reading achievement. *Quarterly Journal of Music Teaching and Learning* 6, no. 1: 56–79.

Lee, J.I. 2004. Component skills involved in sight reading music. PhD diss., Hanover University of Music and Drama, Germany.

———. 2006. The role of inner hearing in sight-reading music as an example of inter-modal perception. *Musikpsychologie* 18: 35–52.

Lehmann, A.C. 1997. The acquisition of expertise in music: Efficiency of deliberate practice as a moderating variable in accounting for sub-expert performance. In *Perception and cognition of music,* ed. I. Deliège and J.A. Sloboda, 161–87. Hove, UK: Psychology Press.

———. 2005. Vomblattspiel und Notenlesen [Sight reading and score reading]. Vol. D/VII/1 of *Allgemeine Musikpsychologie* (Enzyklopädie der Psychologie), eds. T.H. Stoffer and R. Oerter, 877–912. Göttingen, Germany: Hogrefe.

Lehmann, A.C., and K.A. Ericsson. 1993. Sight-reading ability of expert pianists in the context of piano accompanying. *Psychomusicology* 12, no. 2: 182–95.

———. 1996. Performance without preparation: Structure and acquisition of expert sight-reading and accompanying performance. *Psychomusicology* 15, no. 1–2: 1–29.

Lehmann, A.C., J.A. Sloboda, and R.H. Woody. 2007. *Psychology for musicians: Understanding and acquiring the skills.* Oxford, UK: Oxford University Press.

MacMillan, N.A., and C. Creelman. 1991. *Detection theory: A user's guide.* Cambridge, UK: Cambridge University Press.

McPherson, G.E. 1994. Factors and abilities influencing sightreading skill in music. *Journal of Research in Music Education* 42, no. 3: 217–31.

———. 1995. Five aspects of musical performance and their correlates. *Bulletin of the Council for Research in Music Education* 127: 115–21.

McPherson, G.E., M. Bailey, and K.E. Sinclair. 1997. Path analysis of a theoretical model to describe the relationship among five types of musical performance. *Journal of Research in Music Education* 45, no. 1: 103–29.

McPherson, G.E., and A. Gabrielsson. 2002. From sound to sight. In *The science and psychology of music performance. Creative strategies for teaching and learning*, ed. R. Parncutt and G.E. McPherson, 99–115. Oxford, UK: Oxford University Press.

Neubauer, A.C., F.M. Spinath, R. Riemann, P. Borkenau, and A. Angleitner. 2000. Genetic and environmental influences on two measures of speed of information processing and their relation to psychometric intelligence: Evidence from the German observational study of adult twins. *Intelligence* 28, no. 4: 267–89.

Oberauer, K., H.-M. Suess, R. Schulze, O. Wilhelm, and W.W. Wittmann. 2000. Working memory capacity – Facets of a cognitive ability construct. *Personality and Individual Differences* 29, no. 6: 1017–45.

Oswald, W.D., and E. Roth. 1997. *Der Zahlen-Verbindungs-Test [ZVT, Number Combination Test].* Göttingen, Germany: Hogrefe.

Peters, M. 1976. Prolonged practice of a simple motor task by preferred and nonpreferred hands. *Perceptual and Motor Skills* 43, no. 2: 447–50.

Raven, J.C. 2000. *Standard progressive matrices (SPM).* Florence, Italy: Organizazzioni Speciali.

Salis, D.L. 1978. The identification and assessment of cognitive variables associated with reading of advanced music at the piano. PhD diss., University of Pittsburgh. *Dissertation Abstracts International-A* 38, no. 12: 7239–40.

Salthouse, T.A. 2000. Aging and measures of processing speed. *Biological Psychology* 54, no. 1–3: 35–54.

Schleuter, S.L. 1993. The relationship of AMMA scores to sight singing, dictation, and SAT scores of university music majors. *Contributions to Music Education* 20: 57–63.

Sloboda, J.A. 1974. The eye–hand span. An approach to the study of sight reading. *Psychology of Music* 2, no. 2: 4–10.

———. 1984. Experimental studies of music reading: A review. *Music Perception* 2, no. 2: 222–36.

Sloboda, J.A., and D.H.H. Parker. 1985. Immediate recall of melodies. In *Musical structure and cognition,* ed. P. Howell, I. Cross and R. West, 143–67. London: Academic Press.

Steinberg. 2000. *Cubase VST Score 5.1* (Computer software). Hamburg, Germany: Steinberg Ltd.

Thompson, W.B. 1985. Sources of individual differences in music sight-reading skill. Doctoral diss., University of Missouri – Columbia. *Dissertation Abstracts International-B* 47, no. 2: 828.

UNISA. 1995. Vols. 1–8 of *Playing at sight (Piano)*. Pretoria: University of South Africa.

Waters, A.J., E. Townsend, and G. Underwood. 1998. Expertise in musical sight reading: A study of pianists. *British Journal of Educational Psychology* 89: 123–49.

White, B.W. 1960. Recognition of distorted melodies. *American Journal of Psychology* 73: 100–7.

Wolf, T. 1976. A cognitive model of musical sight-reading. *Journal of Psycholinguistic Research* 5, no. 2: 143–71.

Appendix 1.

Warm-up example from the sight-reading task. The pre-recorded pacing voice was played by a violin and played back by speaker. The cue to start playing was indicated by two full bars of clicks before the piece.

Appendix 2.

Warm-up example from the inner hearing test. From top to bottom: Original theme presented to the subjects for 45 seconds (from Beethoven, Six variations on a Swiss song [Woo 66]), lure melody (played to the subjects after scroe disappeared), variation (no. 1) from original composition (played to subjects for reasons of better understanding only in the warm-up examples).