A reliable and valid tool for measuring visual recognition ability with musical notation

Yetta Kwailing Wong¹ · Kelvin F. H. Lui² · Alan C.-N. Wong²

Published online: 1 September 2020
© The Psychonomic Society, Inc. 2020

Abstract

Recognizing musical notation is an important skill to a full participation of Western classical music, but remains a largely under-researched topic in the psychology of music. One plausible reason of such omission is that, in the past, the research field has heavily relied on self-report of music reading ability, which was subjective and highly variable. This paper presents a reliable and valid tool for objectively measuring individual abilities in visual recognition of musical notation. The visual fluency task measures how fast one can accurately recognize a sequence of musical notation at a desired accuracy level using the adaptive psychometric method QUEST. We checked the reliability of this task in over 200 participants in terms of Guttman's λ -2 and Cronbach's alpha. Also, we evaluated the construct validity of this task by considering the convergent validity of this task with multiple external real-world measures of one's musical training background, with numerous experimental measures of visual tendencies of musical notation recognition and with sight-reading performance. Overall, the visual fluency task achieved satisfactory reliability and validity for measuring abilities in recognizing musical notation. This opens the door for characterizing the cognitive mechanisms, development, and individual differences in musical notation recognition, for understanding music learning and music psychology and for understanding of visual perceptual expertise in general.

Keywords Music reading · Perceptual expertise · Psychophysics · Individual differences

Recognizing musical notation is a fundamental ability to learn in formal music training in Western classical music. Efficient reading of music scores is an invaluable asset, if not an essential ability, for any classical musician to develop an extensive repertoire and enjoy collaborative musical activities¹. Fluent

recognition of musical notation is also a prerequisite of music sight-reading, i.e., to play a piece of music when one reads the score for the first time, which is one of the top skills to acquire in formal musical training (Kostka, 1997; Lowder, 1983; McPherson, 1995; McPherson & Gabrielsson, 2002; Wolf, 1976). During music playing, visual processing of the score interacts with many other cognitive processes, including the auditory processing of sounds played by the instrument, motor planning and execution (e.g., fingers, arms, and lip), musical knowledge, memory, and emotional interpretation and expression. A comprehensive understanding of these processes therefore cannot be achieved without understanding visual recognition of musical notation (Sloboda, 2005b)².

In the past few decades, a lot of effort has been devoted to understanding the psychology of music, from the

¹ Note that many great musicians do not read musical notation well, such as the pop musician Paul McCartney and the jazz musician Tony Williams. Also, the majority of the world's musical activity (in past and present) has been performed by musicians who do not read musical notation.

Electronic supplementary material The online version of this article (<https://doi.org/10.3758/s13428-020-01461-w>) contains supplementary material, which is available to authorized users.

✉ Yetta Kwailing Wong
yetta.wong@gmail.com

✉ Alan C.-N. Wong
alanwong@cuhk.edu.hk

¹ Department of Educational Psychology, Faculty of Education, The Chinese University of Hong Kong, Rm 308, Ho Tim Building, Shatin, NT, Hong Kong

² Department of Psychology, The Chinese University of Hong Kong, 344 Sino Building, Shatin, NT, Hong Kong

² In the literature, there are different definitions of the term 'music reading': Some define it as the process of performing music based on musical notation (e.g., Gudmundsdottir, 2010), while others define it as the process of recognizing and understanding musical notation as a form of physical description of sound (similar to what conductors do during score reading; Sloboda, 1978). In this paper, we try to avoid using the term 'music reading' to minimize confusion. Instead, we use the term 'musical notation recognition' to clarify that the focus of our discussion is on the process of 'visual recognition of musical notations'.

understanding of sound, pitch, melody and timbre, the development of musical ability and expertise, to the cognitive neuroscience of music (Deutsch, 1999, 2013; Gordon, 1971; Patel, 2008; Peretz & Zatorre, 2003; Sloboda, 2005a; Spiro, 2003). However, little has been done to investigate the psychology of musical notation recognition and music reading. A review stated that “influential commentators on psychological aspects of music have, in fact, had very little to say about music reading” (p.1, Sloboda, 2005b), illustrating that the past music research has been highly focused on the listeners’ perspective but not also a readers’ perspective.

Previous work in recognizing musical notation relied heavily on self-report of one’s ability or on the assumption that professionally trained musicians are ‘expert music readers’ (Brochard et al., 2004; Brodsky et al., 2003; Furneaux & Land, 1999; Gunter et al., 2003; Kinsler & Carpenter, 1995; Schön et al., 2002; Sergent et al., 1992; Sloboda, 1974; Waters et al., 1997, 1998; Waters & Underwood, 1999). It has been demonstrated that such an assumption is not always valid because the speed of accurately recognizing musical notation is highly varied even within experienced musicians, and the performance of some musicians could actually be more similar to non-musicians than experienced musicians (Figure 2, Y. K. Wong & Gauthier, 2010a). This may perhaps explain why studying musical notation recognition has remained highly difficult. For the study of musical notation recognition to move forward, it is essential to develop a valid and reliable tool to objectively measure one’s ability. Below, we introduce a visual fluency measure that estimates how fast one can accurately recognize musical notation, and discuss the rationale of its design.

Measuring visual fluency of musical notation

The visual fluency task focuses on a particular aspect of visual object recognition - the ability to individuate visually similar objects (music sequences in this case), or to perform subordinate level categorization (Logothetis & Sheinberg, 1996). It uses a sequential matching paradigm to measure how fast one can recognize a sequence of musical notation (Fig. 1), and a similar task has been used to measure the visual perceptual fluency of recognizing English words and Chinese characters (A. C.-N. Wong et al., 2019). Participants first see a target sequence, which is presented briefly on the screen and then disappears. Then, they see two sequences on the screen side-by-side, among which one is identical to the target they just saw, while the other is a highly similar distractor. Their task is to select the identical target among the two options without any time limit.

The key of this task is that the presentation speed of the first target is continuously adjusted based on participants’ performance. In general, the next target would become faster if one

provides a correct answer, and become slower otherwise. The speed adjustment is performed by a computer algorithm to achieve a stable estimate of presentation speed that keeps the participants at about 80% accuracy. In this study, the speed adjustment was done by a Bayesian algorithm called QUEST (Watson & Pelli, 1983), while there are other alternatives that provide similar estimates with different algorithms. For example, another common way to perform such estimation is by using a staircase procedure (e.g., García-Pérez, 1998), which also works well in our separate sample (see [Supplementary Information](#)).

This task has a few important characteristics by design. First, the task pushes the limit of one’s visual abilities in recognizing musical notation. The distractor always differs from the target by only one note, which is shifted upward or downward by one step on the staff. Also, each sequence cannot have any identical pitches before or after such shifts. In this way, performance of this task depends on accurate discrimination of the pitch of the notations, which is exactly what we want to measure. The general contours of the options remain largely identical since pitch contour contributes to the perceived melodic similarity (Prince, 2014). The visual recognition of the rhythmic patterns of the notation is not the focus of this measure because different rhythmic patterns typically involve obvious differences of visual features e.g., stems, tails, or color of the dots (black or white) for which musicians and non-musicians tend to achieve similar performance (Y. K. Wong et al., 2014; Y. K. Wong & Gauthier, 2010a)³. In other words, these salient changes of visual features are unlikely critical to the expertise in musical notation recognition. With a highly similar contour, small differences in pitch and a brief presentation time, ability in extracting the visual details of the music sequences becomes the limiting factor in one’s performance in the task.

Second, the task requires visual judgment and no music knowledge or training to the extent that even complete non-musicians can perform this task. For example, it does not involve any explicit naming of the notation, which tends to be complicated by the various notation naming systems of the learners and extraneous factors that affect performance, including writing fluency and motivation. One big advantage of this is that it enables the investigation of musical notation recognition among individuals of any skill level.

Third, randomly generated sequences are used to further minimize any performance advantage based on musical

³ In these studies, the stimuli included musical notations with differences in both pitch and rhythmic values (e.g., with open circles as in half notes, black circles as in 4th notes, one tail as in 8th notes, and two tails as in 16th notes). In the one-back task that required participants to detect immediate repetition of the images as quickly and as accurately as possible, both studies showed that musicians and non-musicians attained comparable accuracy and response time, suggesting that visually discriminating between notes that differ in *both* pitch and rhythm is unlikely what sets apart individuals with different expertise levels in recognizing musical notations.

Fig. 1 The sequential matching paradigm in the visual fluency task with musical notation

knowledge. Note that we fully recognize the importance of musical knowledge in predicting music sequences during music playing and sight-reading. However, musically meaningful sequences tend to be regular visual patterns that notes are often evenly distributed either between the lines (e.g., the F major chord; Fig. 2) or on the lines. Such visual regularity makes the task easy even for those without any musical training. In contrast, random notation patterns tend to be visually irregular and are hard to predict based on musical knowledge and syntax. In this way, using random sequences allows us to maximally push the visual system to encode detailed visual information during measurement.

Lastly, the task is administered with a computer program which limits the exposure time of the musical notation. Computerization enables a precise and objective measure, which cannot be achieved by any paper-and-pencil tasks commonly adopted in classroom settings, or in previous research that did not use computerized measures (e.g., Sloboda, 1974).

While this visual fluency task has been adopted in several previous studies (Y. K. Wong et al., 2014; Y. K. Wong & Gauthier, 2010a, 2010b, 2012), its reliability and validity remains unclear. The major goal of this paper is to examine the reliability and validity of this tool for measuring one's ability in recognizing musical notation. We assessed the reliability of this task with data from over 200 individuals, and discuss the validity of this measure based on other measures related to music reading experience and ability. Finally, we discuss how this tool may contribute to the understanding of music psychology, music teaching and learning, and research in visual perceptual expertise in general.

Method

Participants

Two hundred and forty-one participants were recruited at the local universities of Hong Kong for various projects studying musical notation recognition in the authors' research group. They included 71 males and 170 females aged between 17 and 60 ($M = 22.13$, $SD = 5.56$). Data from ten participants were discarded as one or more of the visual fluency measures were over 3 SDs away from the average. The remaining participants were 67 males and 164 females aged between 17 and 60 years ($M = 22.06$, $SD = 5.44$). All participants were paid for their participation, and provided informed consent according to the ethics committee of the local universities in which data were collected.

All participants reported their music training background through a questionnaire, including all instruments they played, and the age of onset of the instrumental training and the number of years of learning for each instrument. For each participant, the instrument with the largest number of years he or she has been playing was defined as the major instrument. The questionnaire also asked the number of years of experience one had in music reading and a self-report rating of one's music reading ability from 1 to 10 (1 = Don't read music at all, 10 = Expert in music reading).

Among the 231 remaining participants, 131 had one or more years of formal training in music reading and were defined as experienced readers (Table 1). Their experience in

Fig. 2 An example of a visually regular sequence in F major chord (a) and a relatively irregular sequence that is generated randomly in pitch (b)

Table 1 Musical training background of the participants

I	Music reading experience (years)		Self-report music reading ability		Instrumental training (years)		Onset of instrumental training (years)	
	Experienced	Novice	Experienced	Novice	Experienced	Novice	Experienced	Novice
N	116	100	130	100	125	14	125	14
Mean	10.68	0.01	6.75	1.72	10.83	3.00	7.90	10.29
Median	11.00	0	7.00	1.00	11.00	2.00	7.00	11.00
SD	4.07	0.10	1.97	1.04	3.80	2.97	2.83	3.58
Min	1	0	1	0	1	0.5	4	4
Max	22	0.9	10	6	23	10	17	18

Note. 131 experienced readers and 100 novice readers filled out this questionnaire. Sample sizes are different in each measure because of missing data. The self-report music reading ability was measured on a 10-point rating scale

reading music was 10.7 years on average and the average self-evaluated music reading ability was 6.75 (out of 10). They have been playing their major musical instruments for 10.8 years on average, and the average onset of instrumental training was 7.9 years old. For the 100 novice readers, their experience in reading music was .014 years and the average self-evaluated music reading ability was 1.72 (out of 10). Only 14 of them reported having had musical training. On average they played musical instruments for 3.00 years and the average age of onset of instrumental training was 10.3 years old.

Materials

Each participant performed the fluency task for music sequences of four quarter notes. The pitch of the notation was randomly generated from 11 notes between D4 and G5 (without accidentals) so as to avoid the use of ledger lines. Also, no repetition of pitches was allowed within each target or distractor sequence.

Participants also performed the task with another stimulus category that served as a control category. This enabled one to control for individual differences in perceptual abilities not specifically tied to abilities with musical notation recognition, such as general shape processing, recognizing four visual objects arranged horizontally, recognizing briefly presented stimuli, matching sequentially presented stimuli, etc. The control stimuli were either four-letter strings or four-digit strings. For letters, the following 11 letters were included: b, d, f, g, h, j, k, p, q, t, and y. For digits, all digits from ‘0’ to ‘9’ were used except ‘1’. For the distractor string, the different letter or number was replaced by randomly selecting another letter or number within the stimulus set, and their position within the distractor string was counterbalanced. Contrast for all the stimuli was lowered by about 60% to avoid a ceiling effect. Details of the manipulation of the visual stimulus can be found in Wong and Gauthier (2010a) and Wong & Wong (2016).

Procedure

The visual fluency task was performed on personal computers with Matlab and the Psychtoolbox extension (Brainard, 1997; Pelli, 1997). During each trial, a fixation cross was presented at the center of the screen for 200 milliseconds (ms), followed by a 500 ms pattern as a mask (Fig. 1). Next, a target music sequence appeared for a varied duration, which was replaced by the mask again for 500 ms. Then, two music sequences were presented side-by-side on the computer screen, with one being identical to the target and one being slightly different. The different sequences would have one of the notes shifted upward or downward by one step on the staff, with the shifted note and the shifting direction counterbalanced. Participants were asked to identify the target by pressing either ‘Z’ or ‘M’, which represented the left or the right stimulus, respectively. The current version of this task is already made available (https://osf.io/axzpt/?view_only=35c4d7683645456fa273d58d8b0ebbee) for anyone who are interested in experiencing the task.

The visual fluency threshold of an individual required to achieve 80% accuracy was estimated using QUEST (Watson & Pelli, 1983). The fluency threshold was measured four times, each with 40 trials, for each stimulus category. Visual fluency was calculated by averaging the four estimates. The data supporting the following results are available (https://osf.io/axzpt/?view_only=35c4d7683645456fa273d58d8b0ebbee).

Results

Descriptive results

The descriptive statistics of the music reading fluency were shown in Table 2. As expected, experienced readers had a better performance ($M = 644.1$ ms, $SD = 375.1$) than that of all participants ($M = 914.5$ ms, $SD = 570.4$) and that within novice readers ($M = 1269.9$ ms, $SD = 589.5$). The dispersion

Table 2 Descriptive statistics of the music reading fluency. A smaller number indicates a shorter presentation time required to achieve 80% accuracy and thus better performance

Group	<i>N</i>	Mean (ms)	SD	Min (ms)	Max (ms)
Combined	231	914.53	570.43	162.92	3105.93
Experienced readers	131	644.06	375.14	162.92	1810.77
Novice readers	100	1268.86	589.53	313.33	3105.93

of the fluency of all participants and that within each group are shown in Fig. 3.

Reliability of the visual fluency task

The reliability of the visual fluency task was evaluated using Guttman's λ -2 because the multiple estimates of visual fluency with recognizing musical notation were theoretically equivalent in content. Results showed that the visual fluency task achieved satisfactory reliability among all participants (λ -2 = .772), within experienced readers (λ -2 = .733), and with relatively lower reliability within novice readers (λ -2 = .670; Table 3). We also evaluated the reliability of this task using another popular measure of reliability, Cronbach's alpha, and the results were similar (alpha = .767, .726 and .646 and for all participants, experienced readers, and novice readers, respectively). The reliability of this task estimated with letters (as a control category), with musical notations after partialing out general perceptual fluency (measured with letters), and that with a different staircase algorithm also achieved satisfactory reliability when all participants were considered (see [Supplementary Information](#)).

Validity of the visual fluency task

Is the visual fluency task indeed measuring one's ability in recognizing musical notation? We evaluated the construct validity of this task with the following reasoning. If the construct validity of this task was adequate, it should converge with measures that generally reflected one's amount of experience in musical notation recognition and self-report of one's reading ability (i.e., convergent validity). It should also be associated with behavioral and neural markers of perceptual expertise of musical notation recognition, which were separately measured by other behavioral tasks or neural imaging methods. Finally, since musical notation recognition is fundamental to sight-reading, i.e., to play a piece of music when one reads the score for the first time, visual fluency should also predict sight-reading performance. Below, we report the findings of these aspects of evaluations.

First, performance of the visual fluency task was generally consistent with one's musical training background, based on the data from the same sample contributed to the reliability

measures. For example, experienced readers were in general faster in recognizing musical notation than novice readers (Table 2). Also, among experienced readers, performance in the music reading fluency task could be significantly predicted by the number of years of music reading experience ($r(114) = -.380, p < .001$), the number of years spent playing the major instrument ($r(123) = -.294, p < .001$), and the self-report music reading ability ($r(128) = -.452, p < .001$). Performance in the music reading fluency task, however, was not significantly correlated with the age when training on the major instrument started ($r(123) = .079, p = .382$). These showed that the visual fluency measure had high convergent validity in terms of real-world measures related to one's musical training background. The lack of correlation between music reading fluency and age of onset of music training indicates the lack of strong evidence for any critical period constraining the development of music reading fluency.

Second, individual difference in the visual fluency task predicted various behavioral and neural markers of expertise in recognizing musical notation. The sample of these studies were separately recruited from that reported in this current study, and the visual fluency tasks adopted were highly similar to that reported in the current study, i.e., a sequential matching task with four-note sequences, combined with an adaptive psychometric function for adjusting the presentation duration of the target with QUEST as in the current study, or a staircase procedure that also attains satisfactory reliability (see [Supplementary Information](#)). This allows direct comparison between the findings of the visual fluency task in the previous studies and that of the current study.

At the behavioral level, the visual fluency measure predicted one's perceptual tendency to process a sequence of musical notation as 'a whole object' rather than separate isolated notations ($r(15) = -.566, p = .018$; Y. K. Wong & Gauthier, 2010b). This perceptual tendency, called 'holistic processing', is a common marker of perceptual expertise for many domains of objects including faces (Farah et al., 1998; Young et al., 1987), cars (Gauthier et al., 2003), chessboards (Boggan et al., 2012), words (A. C.-N. Wong et al., 2011) and Chinese characters (A. C.-N. Wong et al., 2012).

Also, it predicted the degree of inappropriate integration of nearby distractors with the to-be-recognized musical stimuli ($r(37) = .339, p = .035$; and $r(37) = .518, p = .0007$ respectively for the two stimulus conditions; Y. K. Wong & Gauthier, 2012). This phenomenon, called 'visual crowding', is regarded a bottleneck in visual object recognition, such as letter and word recognition (Pelli & Tillman, 2008).

Furthermore, it was correlated with the degree in which musical notation recognition was disrupted by the removal of the line junctions between the dot of a musical notation and the five-line staff ($r(65) = -.33, p = .006$ and $r(65) = -.27, p = .026$ for two performance measures respectively; Y. K. Wong & Wong, 2018). The effect of line junction

Fig. 3 The distribution of the visual fluency estimated using QUEST among all participants (a), experienced readers (b) and novice readers (c)

removal on recognition performance is generally observed in many object domains, including letters, words and

common objects (Biederman, 1987; Lanthier et al., 2009; Szwed et al., 2009).

At the neural level, using functional magnetic resonance imaging (fMRI), it was observed that one's visual fluency with musical notation predicted the extent to which brain areas showed specialized neural responses for musical notation compared with control objects and with novice readers ($r(16)$ ranged from $-.477$ to $-.633$, $ps = .005$ to $.045$; Y. K. Wong & Gauthier, 2010a). Also, using electroencephalograph (EEG), it was found that individual visual fluency predicted the specialized neural response for musical notation as early as

Table 3 Reliability of the visual fluency measure for musical notation estimated with QUEST

Participants	<i>N</i>	λ -2	Alpha
All	231	.772	.767
Experienced readers	131	.733	.726
Novice readers	100	.670	.646

40 ms after experienced readers saw a single musical notation ($r(39) = -.32, p = .043$; Y. K. Wong et al., 2014).

Note that these expertise markers did not simply reflect general performance level in these tasks, but indicated specific expert processes because various control conditions were often included. For example, in the visual crowding task (Y. K. Wong & Gauthier, 2012), crowding was measured as the performance difference between the crowded condition and the uncrowded baseline condition. Overall, the visual fluency measure predicted multiple behavioral and neural markers of expertise in recognizing musical notation. These demonstrated that the perceptual expertise measure has high convergent validity in terms of related experimental measures that capture expertise in recognizing musical notation.

Finally, in a large sample of 143 participants (independent of the sample reported here), we used the visual fluency task to quantify individual ability in recognizing musical notation and investigated its relationship with one's sight-reading performance. Results demonstrated that visual fluency is among one of the strongest predictors of music sight-reading, and this effect holds after partialling out the visual fluency for digits as control stimuli (Pu, Wong, & Wong, [under review](#)).

In sum, the visual fluency measure showed convergent validity with external real-world measures based on musical training background. It also showed convergent validity with experimental measures that target on perceptual expertise in recognizing musical notation. Furthermore, it predicted sight-reading performance in which musical notation recognition is critical. All these suggest that the visual fluency task has a satisfactory construct validity for measuring one's ability in recognizing musical notation.

Discussion

This paper introduced the visual fluency measure that estimates how fast one can recognize a sequence of musical notation accurately. Our analyses showed that this measure achieved an adequate reliability among all participants and among experienced readers. Based on external and experimental measures related to musical notation recognition, this measure also demonstrated satisfactory construct validity. With this reliable and valid tool, one can objectively measure individual performance in recognizing musical notation. It opens the door for characterizing the development and individual differences in musical notation recognition, and for understanding how this ability contributes to Western classical music learning and music cognition, which deserves more attention in the literature of music psychology (Sloboda, 2005b). Below we discuss the potential contribution of this tool to the understanding of musical ability as in the case of music sight-reading and teaching and learning classical music,

and to the understanding of visual perceptual expertise in general.

Contribution to music sight-reading research

In the past, research in music sight-reading has largely ignored the role of musical notation recognition (Kopiez & Lee, 2006, 2008). Our results demonstrated that visual fluency with musical notation predicted sight-reading performance (Pu, Wong, & Wong, [under review](#)). Importantly, other predictors of sight-reading are all visual-association abilities, such as one's ability to imagine the sound of musical notation without hearing the actual sound input, and one's ability to generate motor execution accurately and rapidly based on visual input. This corresponds well with recent findings that adding white space between the musical notes enhanced sight-reading performance (Stenberg & Cross, 2019), possibly through reducing visual crowding between musical notations (Y. K. Wong & Gauthier, 2012). These illustrate the importance of understanding the role of visual processes in music research.

Contribution to teaching and learning classical music

In the classroom of teaching and learning classical music, the major assessment requires playing music on one's own instrument, which is the end product of satisfactory mastery of many musical skills. However, in case of dissatisfactory performance, it is highly difficult to pinpoint how much the struggle could be attributed to musical notation recognition per se. This visual fluency measure provides an objective evaluation of one of the skills involved, which allows teachers to isolate this difficulty from many potential candidate factors.

However, it is important to note that this task focuses on visual recognition of the notes, and it does not capture the conceptual understanding of the notations. For example, this task does not reflect whether the time value represented by each note has been grasped accurately. Since no clef signs were presented, this task does not inform whether the participant can map the visual notation onto the correct note to be performed on his or her musical instrument. Besides, this visual fluency measure focused on the eleven musical notations within the five-line staff (from D4 to G5) without including ledger lines or accidentals. While these were not included such that the difficulty of the task would not be too high and therefore the reliability of the task would be more satisfactory, they are often anecdotally reported to be the more challenging parts during musical notation recognition. Future studies should consider incorporating them into the design of specific measurements for more advanced musicians.

The visual fluency measure also provides an important link and intervention for individuals who encounter difficulties in recognizing musical notation. Anecdotally, musical notation recognition is one of the major reasons why students quit classical music training (based on our personal communications with music teachers). However, music teachers often report that, even they excel in fluent reading of musical notation, they do not know ‘what to say to the students’ who do not (Sloboda, 2005b; Wolf, 1976). Importantly, the rich literature of perceptual expertise has demonstrated that a few hours of perceptual training in the laboratory can lead to substantial improvement in perceptual skills in recognizing a particular category of objects, which is often accompanied with large-scale changes in brain activities for those objects (Gauthier et al., 1999; A. C.-N. Wong et al., 2009; Y. K. Wong et al., 2012). Such perceptual training works for musical notation recognition, since previous work has showed that eight hours of laboratory training can dramatically improve the visual fluency with short note sequences such that performance of the trained individual approaches the level of real-world experienced readers by the end of the training (Y. K. Wong & Wong, 2016).

Although this particular study used four-to-five-note sequences as the training stimuli, the training of music sequences can be tailor-made to cater for the need of different readers. For readers who often read one-staff lines of music, they can focus on expanding the horizontal visual span by gradually increasing the number of notations exposed in a glance. For pianists who need to read two staves at once, they can focus on expanding both the horizontal and vertical visual span such that they can capture the notes on both staves without moving their eyes. Hence this training can potentially be widely useful for assisting students in music learning to overcome the obstacles in visual recognition. Notably, this is completely computerized, which makes training more assessable to those who do not have the resources for professional teachers.

Given the many potential advantages of the visual fluency measure in music research and classical music learning, future studies should work on more accessible versions of this measurement, and develop user-friendly versions of related perceptual training for students in need.

Contribution to research in visual perceptual expertise

To understand individual differences in visual perceptual expertise and its development, it is important to have a measurement tool that quantifies the skill level of an individual. Perceptual experts can efficiently and effortlessly differentiate between similar visual objects within their expertise domains (Gauthier et al., 2009). This is exactly what the visual fluency is designed to measure – the

ability to individuate visually similar instances in an object category. The sequential matching task with adaptive presentation speed of the target stimulus is not specific to measuring musical notation recognition, but can be applied to measuring perceptual expertise in other visual domains, such as the recognition of Chinese characters and English words (A. C.-N. Wong et al., 2019). The satisfactory reliability and validity of this task demonstrates the potential of this task to help quantify skill levels in different domains of visual perceptual expertise. This would enable one to track the developmental trajectory and individual differences of perceptual expertise, and to compare and contrast perceptual expertise in different domains.

Open Practices Statement The data and the program (written in Matlab) are available at https://osf.io/axzpt/?view_only=35c4d7683645456fa273d58d8b0ebbee.

References

- Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, 94, 115–147.
- Boggan, A. L., Bartlett, J. C., & Krawczyk, D. C. (2012). Chess masters show a hallmark of face processing with chess. *Journal of Experimental Psychology: General*, 141(1), 37–42. <https://doi.org/10.1037/a0024236>
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433–436.
- Brochard, R., Dufour, A., & Després, O. (2004). Effect of musical expertise on visuospatial abilities: evidence from reaction times and mental imagery. *Brain and Cognition*, 54(2), 103–109. [https://doi.org/10.1016/S0278-2626\(03\)00264-1](https://doi.org/10.1016/S0278-2626(03)00264-1)
- Brodsky, W., Henik, A., Rubinstein, B. S., & Zorman, M. (2003). Auditory imagery from musical notation in expert musicians. *Perception & Psychophysics*, 65(4), 602–612. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=AbstractPlus&list_uids=12812282
- Deutsch, D. (1999). *The Psychology of Music*. Academic Press.
- Deutsch, D. (2013). *The Psychology of Music*. In *The Psychology of Music* (3rd ed.). Elsevier Science & Technology. <https://doi.org/10.1016/C2009-0-62532-0>
- Farah, M.-J., Wilson, K.-D., Drain, M., & Tanaka, J.-N. (1998). What is “special” about face perception? *Psychological Review*, 105(3), 482–498.
- Furneaux, S., & Land, M. F. (1999). The effects of skill on the eye-hand span during musical sight-reading. *Proc Biol Sci*, 266(1436), 2435–2440. <https://doi.org/10.1098/rspb.1999.0943>
- García-Pérez, M. A. (1998). Forced-choice staircases with fixed step sizes: Asymptotic and small-sample properties. *Vision Research*, 38, 1861–1881. [https://doi.org/10.1016/S0042-6989\(97\)00340-4](https://doi.org/10.1016/S0042-6989(97)00340-4)
- Gauthier, I., Curran, T., Curby, K. M., & Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nat Neurosci*, 6(4), 428–432. <https://doi.org/10.1038/nn1029>
- Gauthier, I., Tarr, M., & Bub, D. (2009). Perceptual expertise. In *Perceptual Expertise: Bridging Brain and Behavior*. Oxford

- University Press. <https://doi.org/10.1093/acprof:oso/9780195309607.001.0001>
- Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform “face area” increases with expertise in recognizing novel objects. *Nature Neuroscience*, 2(6), 568–573.
- Gordon, E. (1971). *The Psychology of Music Teaching*. Prentice Hall.
- Gudmundsdottir, H. R. (2010). Advances in music-reading research. *Music Education Research*, 12(4), 331–338. <https://doi.org/10.1080/14613808.2010.504809>
- Gunter, T. C., Schmidt, B. H., & Besson, M. (2003). Let’s face the music: a behavioral and electrophysiological exploration of score reading. *Psychophysiology*, 40(5), 742–751. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=AbstractPlus&list_uids=14696727
- Kinsler, V., & Carpenter, R. H. (1995). Saccadic eye movements while reading music. *Vision Research*, 35(10), 1447–1458. <http://www.ncbi.nlm.nih.gov/pubmed/7645273>
- 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. <https://doi.org/10.1080/14613800701871363>
- Kostka, M. J. (1997). Effects of self-assessment and successive approximations on “knowing” and “valuing” selected keyboard skills. *Journal of Research in Music Education*, 45(2), 273–281.
- Lanthier, S. N., Risko, E. F., Stolz, J. A., & Besner, D. (2009). Not all visual features are created equal: Early processing in letter and word recognition. *Psychonomic Bulletin & Review*, 16(1), 67–73. <https://doi.org/10.3758/PBR.16.1.67>
- Logothetis, N. K., & Sheinberg, D. L. (1996). Visual object recognition. *Annual Review of Neuroscience*, 19, 577–621.
- Lowder, J. E. (1983). Evaluation of keyboard skills required in college class piano programs. *Contributions to Music Education*, 10, 33–38.
- McPherson, G. E. (1995). Five aspects of musical performance and their correlates. *Bulletin of the Council for Research in Music Education*, 115–121.
- McPherson, G. E., & Gabrielsson, A. (2002). From sound to sight. In R. Parncutt & G. McPherson (Eds.), *The science and psychology of music performance: Creative strategies for teaching and learning* (pp. 99–115). Oxford University Press.
- Patel, A. D. (2008). *Music, Language, And The Brain*. Oxford University Press.
- Pelli, D. G. (1997). The videotoolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, 10, 437–442.
- Pelli, D. G., & Tillman, K. (2008). The uncrowded window of object recognition. *Nat Neurosci*, 1129–1135. <https://doi.org/10.1038/nn.2187>
- Peretz, I., & Zatorre, R. J. (2003). *The cognitive neuroscience of music*. Oxford University Press.
- Prince, J. B. (2014). Contributions of pitch contour, tonality, rhythm, and meter to melodic similarity. *Journal of Experimental Psychology: Human Perception and Performance*, 40(6), 2319–2337. <https://doi.org/10.1037/a0038010>
- Pu, F., Wong, A. C.-N., & Wong, Y. K. (under review). *The role of visual shape processing and visual association abilities in music sight-reading*.
- Schön, D., Anton, J. L., Roth, M., & Besson, M. (2002). An fMRI study of music sight-reading. *Neuroreport*, 13(17), 2285–2289. <https://doi.org/10.1097/01.wnr.0000044224.79663.f5>
- Sergent, J., Zuck, E., Terriah, S., & MacDonald, B. (1992). Distributed neural network underlying musical sight-reading and keyboard performance. *Science*, 257(5066), 106–109. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=AbstractPlus&list_uids=1621084
- Sloboda, J. (1974). The eye-hand span - An approach to the study of sight reading. *Psychology of Music*, 2(2), 4–10.
- Sloboda, J. (1978). The psychology of music reading. *Psychology of Music*, 6, 3–20.
- Sloboda, J. (2005a). *Exploring the musical mind: Cognition, emotion, ability, function*. Oxford University Press.
- Sloboda, J. (2005b). The Psychology of Music Reading. In J. Sloboda (Ed.), *Exploring the Musical Mind: Cognition, emotion, ability, function* (pp. 1–27). Oxford University Press.
- Spiro, J. (2003). Music and the brain. *Nat Neurosci*, 6(7), 661.
- Stenberg, A., & Cross, I. (2019). White spaces, music notation and the facilitation of sight-reading. *Scientific Reports*, 9(1), 5299. <https://doi.org/10.1038/s41598-019-41445-1>
- Szwed, M., Cohen, L., Qiao, E., & Dehaene, S. (2009). The role of invariant line junctions in object and visual word recognition. *Vision Research*, 49(7), 718–725. <https://doi.org/10.1016/j.visres.2009.01.003>
- Waters, A. J., Townsend, E., & Underwood, G. (1998). Expertise in musical sight reading: A study of pianists. *British Journal of Psychology*, 89, 123–149.
- Waters, A. J., & Underwood, G. (1999). Processing pitch and temporal structures in music reading: Independent or interactive processing mechanisms? *European Journal Of Cognitive Psychology*, 11(4), 531–553.
- Waters, A. J., Underwood, G., & Findlay, J. M. (1997). Studying expertise in music reading: use of a pattern-matching paradigm. *Perception & Psychophysics*, 59(4), 477–88.
- Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Percept Psychophys*, 33, 113–120.
- Wolf, T. (1976). A cognitive model of musical sight-reading. *Journal of Psycholinguistic Research*, 5(2), 143–171. <https://doi.org/10.1007/BF01067255>
- Wong, A. C.-N., Bukach, C. M., Hsiao, J., Greenspon, E., Ahern, E., Duan, Y., & Lui, K. F. H. (2012). Holistic processing as a hallmark of perceptual expertise for nonface categories including Chinese characters. *Journal of Vision*, 12, 7. <https://doi.org/10.1167/12.13.7>
- Wong, A. C.-N., Bukach, C. M., Yuen, C., Yang, L., Leung, S., & Greenspon, E. (2011). Holistic processing of words modulated by reading experience. *PLoS ONE*, 6. <https://doi.org/10.1371/journal.pone.0020753>
- Wong, A. C.-N., Palmeri, T., Rogers, B. P., Gore, J. C., & Gauthier, I. (2009). Beyond shape: How you learn about objects affects how they are represented in visual cortex. *PLoS ONE*, 4(12), e8405.
- Wong, A. C.-N., Wong, Y. K., Lui, K. F. H., Ng, T. Y. K., & Ngan, V. S. H. (2019). Sensitivity to configural information and expertise in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 45(1), 82–99. <https://doi.org/10.1037/xhp0000590>
- Wong, Y. K., Folstein, J. R., Gauthier, I. (2012). The nature of experience determines object representations in the visual system. *Journal of Experimental Psychology: General* (Vol. 141, pp. 682–698). <https://doi.org/10.1037/a0027822>
- Wong, Y. K., & Gauthier, I. (2010a). A multimodal neural network recruited by expertise with musical notation. *Journal of Cognitive Neuroscience*, 22(4), 695–713. <https://doi.org/10.1162/jocn.2009.21229>
- Wong, Y. K., & Gauthier, I. (2010b). Holistic processing of musical notation: Dissociating failures of selective attention in experts and novices. *Cognitive, Affective & Behavioral Neuroscience*, 10, 541–551. <https://doi.org/10.3758/CABN.10.4.541>
- Wong, Y. K., Gauthier, I. (2012). Music-reading expertise alters visual spatial resolution for musical notation. *Psychonomic Bulletin & Review* (Vol. 19, pp. 594–600). <https://doi.org/10.3758/s13423-012-0242-x>
- Wong, Y. K., Peng, C., Fratus, K. N., Woodman, G. F., & Gauthier, I. (2014). Perceptual expertise and top-down expectation of musical

- notation engages the primary visual cortex. *Journal of Cognitive Neuroscience*, 26(8), 1629–1643. https://doi.org/10.1162/jocn_a_00616
- Wong, Y. K., & Wong, A. C.-N. (2016). Music-reading training alleviates crowding with musical notation. *Journal of Vision*, 16(8), 15. <https://doi.org/10.1167/16.8.15>
- Wong, Y. K., & Wong, A. C.-N. (2018). The role of line junctions in object recognition: The case of reading musical notation. *Psychonomic Bulletin & Review*, 25(4), 1373–1380. <https://doi.org/10.3758/s13423-018-1483-0>
- Young, A. W., Hellawell, D., & Hay, D. (1987). Configural information in face perception. *Perception*, 10, 747–759.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.