Distractor familiarity reveals the importance of configural information in musical notation

Ting-Yun Chang¹ · Isabel Gauthier¹

Published online: 19 August 2019
© The Psychonomic Society, Inc. 2019

Abstract

The study of perceptual expertise in a visual domain requires the definition of boundaries for the objects that are part of the domain in question. Unlike other well-studied domains, such as faces or words, the domain of musical notation has been lacking in efforts to identify critical features that define the objects of music reading. In the present study, we took advantage of the distractor familiarity effect in visual search. We asked participants to search for a prespecified target note among familiar/unfamiliar distractor notes when two features of musical notation, dot–stem configuration (the way of connecting the dot and the stem of a note) and connectedness (whether or not the dot and the stem of a note were connected), were manipulated. A participant's level of music-reading expertise predicted the magnitude of the distractor familiarity effect only when the dot–stem configuration was diagnostic for the search. Connectedness did not induce a distractor familiarity effect, regardless of its diagnosticity. Dot–stem configuration is a defining feature of music notes, helping to characterize the boundaries of the domain of music-reading expertise. This work has also improved on the tasks used to quantify expertise in reading musical notation.

Keywords Visual expertise · Object recognition · Music · Visual search

Experience in a perceptual domain—for instance, bird watching, reading radiographs, or learning musical notation—can result in impressive levels of performance, but this expertise tends to be specific. Not only does experience with birds not make one a better music reader, but even within the same domain, expertise effects usually are fairly narrow. For instance, although modern-car experts use holistic processing to recognize modern cars, this strategy does not generalize to antique cars (Bukach, Phillips, & Gauthier, 2010). While we know that the features of modern and antique cars differ in some ways, it is not trivial to specify what properties of these complex objects limit the transfer of expertise. Here we addressed a similar question, but in a domain in which objects of expertise are much more simple—the domain of musical notation. As musicians learn to read musical notation, they learn to interpret simple markings such as dots with stems, positioned systematically on arrays of lines (i.e., the staff), which together represent sequences of pitches and the durations of the notes that should be played. Although

musical scores can often be quite complex, the basic units are strikingly simple. It may therefore be easier to ask what the critical features of perceptual expertise for musical notation are than in domains in which the objects are more complex, like birds or cars.

Specifying what aspects of a stimulus define the boundaries of a domain is an important step in the study of expertise, for both theoretical and methodological reasons. On the theoretical side, expertise originates from extensive experience performing one or several tasks with an object category, and it is by nature understood to be domain-specific (e.g., a chess expert is not assumed to be particularly skilled at reading radiographs). Although this seems to be obvious, it is difficult to characterize the exact boundaries of a specific domain. For instance, expertise in reading chest radiographs may or may not extend to reading other kinds of radiographs. Therefore, it is impossible to invoke the idea of a domain without making some inferences about what its boundaries are. For example, in the case of face expertise, studies have suggested that the domain extends to stimuli as different as cartoon characters and faces composed of other items, like fruits (Moscovitch, Winocur, & Behrmann, 1997). This suggests that a critical feature of the domain is shared across all these stimuli, likely the configuration of the parts. When this configuration is disrupted—for instance, when the top and bottom halves of

✉ Ting-Yun Chang
ting-yun.chang@vanderbilt.edu

¹ Vanderbilt University, Nashville, TN, USA

a face are misaligned—mechanisms thought to be specific to face expertise (holistic processing) are not engaged, demonstrating the limits of this generalization (e.g., Hole, 1994; Richler & Gauthier, 2014; Richler, Gauthier, Wenger, & Palmeri, 2008a; Richler, Tanaka, Brown, & Gauthier, 2008b; Tanaka & Sengco, 1997). On the methodological side, the study of expertise is facilitated by the availability of control stimuli that are as similar as possible to those that fall in the domain of interest, yet that do not engage expert mechanisms. For example, in the study of face expertise, inverted faces are often used for this purpose (e.g., McKone et al., 2013). Because each type of expertise combines experience with distinct stimuli and tasks, this work of defining the boundaries of a domain needs to be conducted separately for each domain.

Although the study of expertise in musical notation is relatively recent, a number of interesting behavioral and neural effects of expertise have been identified in this domain. Experience reading music facilitates the encoding of short sequences of notes, such that music-reading experts are almost as efficient with notes as with letters, whereas music-reading novices tend to be much slower with notes than with letters (Wong & Gauthier, 2010a). Interestingly, music-reading experience results in expertise effects that can be obtained for notation at a scale smaller than sequences of notes. The presentation of a single note on staff lines, as compared to single letters or mathematical symbols, is sufficient to elicit an early visual cortex response that correlates with expertise at reading musical notation (Wong & Gauthier, 2010b). Subsequent event-related potential work further revealed that such neural effects of expertise with notes occur as early as 40 ms post-stimulus-presentation, even when a single note is shown without staff lines (Wong, Peng, Fratus, Woodman, & Gauthier, 2014). Behaviorally, music-reading experts excel at very simple judgments on very simple music-relevant stimuli. For instance, experts are better than novices at judging whether a dot is on or off a line on an array of horizontal lines, whereas no such advantage is observed when the lines are vertical (Wong et al., 2014). Despite this growing body of knowledge, along with evidence that effects of expertise in the music-reading domain can be found with simple music-like stimuli, few studies have addressed what the features are that define musical notation.

The features that define musical notation might be roughly grouped into two categories. One type of feature includes those that are critical for reading music, because they contain information diagnostic about the pitch or duration of notes. One example is the contact between notes and staff, or line junctions; fluency in matching short music sequences in which the line junctions are removed is diminished proportionally with the level of music-reading expertise (Wong & Wong, 2018). Music-reading expertise is also sensitive to line orientation. In music reading, notes are always presented on a horizontal staff, and the position of a note on a specific staff

line defines its pitch; music-reading experts outperform novices in judging if a dot is on or off a line when these lines are horizontal, but not when they are vertical (Wong et al., 2014). In contrast, another type of feature is one that appears systematically in musical notation but does not have diagnostic value for the task of music reading. For example, although attaching a line to a note may be meaningful in changing the duration of a note (i.e., a whole note with a line added to it becomes a half note, which is half of its duration), whether the line is attached to the right or the left of the note does not provide information about its pitch or duration. Conventionally, the rule is that lines attached on the right of the note stick up and lines attached on the left of the note go down. This configuration does not provide information about the note identity in the context of music reading, but it may nonetheless define whether an object is a valid part of musical notation. In other domains, there are examples of such nondiagnostic features that do not affect object identity but that, if modified, reduce expertise effects. For instance, expert readers recognize letters less efficiently when print is not regular in font (Gauthier, Wong, Hayward, & Cheung, 2006). Recognizing upside-down faces is more difficult than recognizing upright faces, an effect that is typically larger than for nonface objects (the face inversion effect; Yin, 1969). In principle, the same feature—orientation, configuration, or line junctions—could be more or less important for different domains of expertise, depending on our experience and the regularities of the domain.

In the present study, we considered two features that are nondiagnostic for pitch or duration but that are prevalent and systematically used in written music: dot–stem configuration and connectedness. We asked whether they have become an integral part of what makes musical notation a domain of perceptual expertise. Here we defined dot–stem configuration as the way of connecting the dot and the stem of a note. In staff notation, notes are placed on a staff that consists of five parallel horizontal lines. Typically, stems of notes below the third line point up on the right of the notes, whereas the stems of notes on or above the third line trail down from the left of the notes (Fig. 1). The other feature, connectedness, is defined by whether or not the dot and the stem are attached. Notes are not always coupled with stems, but when there is a stem, it is always connected to the dot. Connectedness may seem in some way similar to the concept of line junction reported in Wong and Wong (2018); however, connectedness does not modify pitch or duration (in that nonconnected stems are never used). Note that we focused on single notes that are not presented on a staff, because using single notes shown without staff allowed us to test whether these nondiagnostic features are sufficient to elicit a music perception in experts at the scale of an individual note, without the context of music reading.

We took advantage of the distractor familiarity effect found in visual search (e.g., Mruczek & Sheinberg, 2005; Reicher, Snyder, & Richards, 1976; Richards & Reicher, 1978; Wang,

Fig. 1 Examples of (a) notes that are made in the correct dot–stem configuration, and (b) notes that are not. When a note has a stem, the stem and the dot are always connected as shown in the top of panel a

Cavanagh, & Green, 1994; Wolfe & Horowitz, 2004) to ask whether dot–stem configuration and connectedness are critical features for music-reading experts. When participants search for a prespecified target among distractors, search is generally found to be more efficient as the distractors become familiar for participants. For instance, when German and Slavic readers search for an *N* (from the Latin alphabet) among mirror-*N*s (from the Cyrillic alphabet), or vice versa, only Slavic readers, who are familiar with both letters, searched efficiently in both conditions, whereas German readers were more efficient searching for a mirror-*N* among *N*s (Malinowski & Hübner, 2001). Similar effects are found with *Z*s and mirror-*Z*s in English readers (Wang et al., 1994), and with Chinese characters in Chinese readers and non-Chinese readers (Shen & Reingold, 2001). In the present study, we adopted a note search task in which participants searched for a prespecified target note among distractor notes, while dot–stem configuration and connectedness were manipulated to be either the same as or different from the form in which they appear in real music-reading contexts. We predicted that, if these features, because of their ubiquity in musical notation, are important for defining the boundaries of what qualifies as musical notation, those who can read music should find it easier to detect the target note when the distractor notes show these features in their familiar versions.

To preview the findings, in Experiment 1 we asked participants to look for a target note among distractor notes with either valid or invalid dot–stem configuration. We also wanted to know, if participants were sensitive to this feature, whether the

effect depended on connectedness, so in each block we varied the connectedness validity for all notes in each display. We found that expertise in music reading predicted the magnitude of the distractor familiarity effect driven by dot–stem configuration, no matter whether or not the parts of the notes were connected. In Experiment 2 we tested whether connectedness itself was able to drive a distractor familiarity effect, by asking participants to look for a target note among distractor notes with either valid or invalid connectedness. We also varied in each block the dot–stem configuration validity for all notes in each display, to see whether this factor would moderate a possible sensitivity to connectedness. The results were consistent with those of Experiment 1 in revealing no sensitivity to connectedness.

Experiment 1

Method

Participants

Before running the study, we estimated the effect size for the distractor familiarity effect of expertise as $r = .40$, on the basis of prior work looking at effects of expertise in a visual search task (McGugin, McKeef, Tong, & Gauthier, 2011). To detect this effect with 75% power at an alpha level of .05, we needed 39 participants. We considered that this level of power would be a reasonable compromise, given the challenge of recruiting participants who varied in their music-reading ability. Forty people (18 males and 22 females) from the Vanderbilt community participated in the study for either course credits or monetary reward. Their ages ranged from 18 to 40 ($M = 23.47$ years, $SD = 5.85$). All of them had normal or corrected-to-normal vision. Although we did not use self-report to determine expertise, we recruited individuals who varied in their self-reported experience in reading music (13 participants, seven of them female, reported extensive experience).

Design, stimuli, and procedure

Sequential matching task Following previous work (Wong & Gauthier, 2010a, 2010b, 2012; Wong et al., 2014; Wong & Wong, 2018), we used a sequential-matching task to quantify participants' perceptual fluency in reading short sequences of musical notation. The same task with digits was used as a control for general processes not specific to music reading. On each trial, a mask was shown for 500 ms, followed by a first (reference) sequence of four notes on a staff, which was presented for a duration determined on the basis of performance on the previous trial. After the first sequence, another 500-ms mask was shown, followed by two test sequences appearing side by side. Participants were instructed to indicate which of these two sequences was identical to the reference

sequence via keypress as accurately as possible. The test sequences remained on screen until a response was made (see Figs. 2a and 2b for examples).

To estimate the exposure duration threshold, the exposure duration for the reference sequence changed on the basis of a 3-up–1-down adaptive method to achieve a threshold of 80%

Fig. 2 Example trials in blocks for matching (a) notes, (b) digits, and (c) shape sequences (the shapes were only used in Exp. 2, as an additional control)

accuracy. This exposure duration started with 700 ms and was shortened by 30% after three consecutive successful trials, or increased by 35% otherwise. Change in the direction of the exposure duration was recorded as a reversal. Estimation terminated when there were five consecutive reversals and at least ten reversals in total, meaning the performance may have reached an asymptote. Therefore, the number of trials in each estimate for each participant varied depending on the steadiness of their performance; however, there were a maximum of 200 trials for each estimate.

The reference music note sequences consisted of four notes on a five-line staff. The notes ranged from the note just below the first line (i.e., a D note in the treble clef) to the note just above the fifth line (i.e., a G note in the treble clef). Distractor music note sequences had one of the four notes shifted by one step (e.g., a note on a line was moved to an adjacent space, or a note on a space was moved to an adjacent line). Shifts were randomly chosen from among the four notes, with upward/downward shifts counterbalanced. As a control for individual differences in non-music-specific processes, we also measured exposure duration thresholds for matching digit sequences using the same procedure. The reference digit sequences were six-digit sequences with no repeated digit (prior work had determined that fewer digits led to ceiling effects; Wong & Wong, 2018). One of the six digits in the reference digit sequence was randomly selected and replaced with a different digit to make the corresponding distractor digit sequences. Still, none of the six digits was ever repeated in a distractor digit sequence. There were two estimation series for music note sequences, and two for digit sequences. Final exposure duration thresholds for matching these sequences were derived by averaging the two series' estimates.

Music note search task Following previous work (Malinowski & Hübner, 2001), we used a search task for prespecified target notes to measure distractor familiarity effects. In the beginning of each block, participants were given a target note that was to be the relevant target for a block of five practice trials and 80 test trials. There were eight blocks throughout the whole experiment, presented in a random order across participants, each using a completely different target (see below for details).

On each trial (see Fig. 3), participants fixated for 500 ms and then saw a display of ten notes—they had to decide whether the target note relevant to the current block was present or absent, as accurately and quickly as possible. There were equal numbers of target-present and target-absent trials. The search display remained on screen until a response was made. The intertrial interval was 1,000 ms. We used a fixed set size of 10 because our main goal was not to determine whether search strategies were qualitatively different (e.g., parallel or serial) across different levels of music-reading expertise. Instead, on the basis of previous work on distractor familiarity (e.g., Malinowski & Hübner, 2001), a set size of 10 was

judged to be large enough to detect an effect of expertise across conditions in correct reaction times (e.g., faster search for unfamiliar over familiar targets, as a function of expertise).

The properties of the notes were set as follows. The eight target notes for each of the eight blocks of trials were composed of the combination of three properties: rightward or leftward location of the line relative to the dot, orientation (whether the line is above or below the dot), and connectedness. The rightward/leftward location of the line and the orientation of the note established the validity of the dot–stem configuration (see Fig. 1).

As a rule, for all target-present displays, the target was either the only note with valid (or invalid) dot–stem configuration—all the distractor notes had a different validity on this dimension. Therefore, dot–stem configuration validity was a diagnostic search feature. If dot–stem configuration is a defining feature of musical notation expertise, then we would predict that music-reading experts would be particularly efficient in searching for the targets in arrays of distractors with valid dot–stem configurations. Connectedness, in contrast, was not a diagnostic search feature, because on all trials, all notes had the same value on this feature. Either all notes (target and distractors) were connected, or they were not, and this remained the same within each block. This was done so that we could examine whether connectedness would modulate how dot–stem configuration was perceived.

Results

Perceptual fluency The two exposure duration threshold estimates were positively correlated with each other for both musical sequences ($r = .56$, $p < .001$, 95% CI = [.30, .74]) and digit sequences ($r = .42$, $p < .01$, 95% CI = [.13, .65]), indicating the reliability of the estimation procedure. The Spearman–Brown correction (Brown, 1910; Kaplan & Saccuzzo, 2017; Spearman, 1910) was applied to correct the r values for both correlations, to verify the reliability of the estimation procedure (music note sequences, $r' = .72$; digit sequences, $r' = .59$). The two threshold estimates for each category were averaged, except for two participants who had one of their threshold estimates exceed 3.5 standard deviations above the mean. In those two cases, the other estimate alone was used. The mean note threshold was 900 ms, with a standard deviation of 560 ms, and the mean digit threshold was 392 ms, with a standard deviation of 200 ms. As expected, we found more variability in the fluency judgments for notes than with digits, since we explicitly recruited participants who varied in their expertise reading music. To verify that the note threshold captured variance specifically tied to musical notation, we established that the first note threshold predicted the second note threshold, even after controlling for variance in digit fluency (the average of the two digit thresholds): $r = .77$, $p < .001$, 95% CI = [.60, .87] (see Fig. 4). For the subsequent analyses, we defined an expertise index as the inverse of the

Fig. 3 Trial procedure of the music-note search task used in Experiments 1 and 2

mean note threshold, controlling for the inverse of the mean digit threshold (the inverse transformation was performed so that higher values would represent better performance).

Music note search task The accuracy in the task was high ($M = .93$, $SD = .26$). Correct reaction times between 200 and 8,000 ms were extracted for further analysis (a total of 7.85% incorrect [7.50%] and outlier [0.35%] data points were excluded from analysis). This range was chosen because 8,000 ms fell outside four standard deviations from the mean reaction times, and we assumed that responses faster than 200 ms were too fast to be meaningful. Log-transformed reaction times were used, to correct for the positive skewness of the original distribution.

We first tested whether music-reading fluency has any relationship with performance searching for a familiar versus an unfamiliar target note, as defined by the valid/invalid dot-stem configuration of the target note. We hypothesized that, if music-

reading experts are indeed sensitive to dot-stem configuration, then the more skilled experts should show the larger distractor familiarity effect—that is, the reaction time difference between conditions of searching for an invalid configural target among valid configural distractors versus conditions of searching for a valid configural target among invalid configural distractors. Correlation analysis confirmed this hypothesis: The expertise index was positively correlated with the magnitude of the distractor familiarity effect ($r = .53$, $p < .001$, 95% CI = [.26, .72]; Fig. 5a). To ensure that this effect was not driven by the presence of outliers, we used the skipped correlation, a robust measure that protects against the influence of bivariate outliers (Pernet, Wilcox, & Rousselet, 2013; Rousseeuw, 1984; Rousseeuw & Van Driessen, 1999; Verboten & Hubert, 2005). The results confirmed the positive correlation between an individual's expertise and the distractor familiarity effect ($r_{\text{skipped}} = .57$, $p < .001$, 95% CI = [.30, .76]). In addition, the level of expertise only correlated with speed when searching for invalid

Fig. 4 Scatter plots of the first and the second exposure duration estimates for note sequences, after controlling for the control conditions, in Experiments 1 and 2 (values are residual scores)

Fig. 5 Scatter plots of (a) the distractor familiarity effect on reaction times, (b) reaction times for detecting an invalid target, and (c) reaction times for detecting a valid target, as a function of the expertise index in

Experiment 1, and one example of each condition. Note that the targets are marked by a red square, which did not appear in the actual experimental sessions

targets ($r = -.48, p < .01, 95\% \text{ CI} = [-.69, -.20]$) (Fig. 5b), but not when searching for valid targets ($r = -.18, p = .27, 95\% \text{ CI} = [-.14, .47]$) (Fig. 5c). Together, these results suggest that, the more fluently one reads music, the faster one detects an invalid configural target among valid configural distractors, whereas when searching for a valid configural target among invalid configural distractors, all participants performed the same, regardless of their music-reading expertise. We also considered the same correlations with error rate, but no significant relationship was found ($r_s < .18, p_s > .27$).

We further tested whether this distractor familiarity effect of validity was modulated by the connectedness between the dot and stem (see Fig. 6). We hypothesized that, since the dot and stem of a note are always connected in formal musical notation, the observed distractor familiarity effect induced from dot–stem configuration might be diminished when all the notes had their parts not connected versus when the notes remained intact. Correlation analyses showed that the expertise index was

moderately correlated with the distractor familiarity effect, in conditions in which the notes were either connected ($r = .42, p < .01, 95\% \text{ CI} = [.13, .65]$) or not connected ($r = .49, p < .01, 95\% \text{ CI} = [.21, .70]$). A z test after Fisher's r -to- z transformation (Fisher, 1921) showed that the two correlations were not significantly different ($z = 0.38, p = .70$). Similarly, the expertise index only predicted the reaction times for detecting an invalid configural target among valid configural distractors ($r_s > .45, p_s < .01$), yet not for detecting a valid configural target among invalid configural distractors ($r_s < .19, p_s > .24$), regardless of the connectedness of the notes. The distractor familiarity effect was robust and was not affected by connectedness.

Discussion

In Experiment 1 we examined whether dot–stem configuration is an important basic feature of a music note, and if so, whether connectedness would modulate its effect. The results of

Fig. 6 Scatter plots of the reaction times for detecting an invalid target under conditions in which all notes were (a) connected or (b) not connected, as a function of the expertise index in Experiment 1, and

one example of each condition. Note that the targets are marked by a red square, which did not appear in the actual experimental sessions

Experiment 1 showed that perceptual fluency in reading music notes predicted the magnitude of the distractor familiarity effect induced from the validity of the notes' dot–stem configuration; a larger effect on reaction times was found in participants that were more fluent in reading notes. Specifically, the effect was due to better performance in search for an invalid target note among valid distractor notes for music-reading experts. Reaction times for searching for a valid target among invalid distractors did not relate to music-reading expertise. Music-reading expertise is thus related to sensitivity to dot–stem configuration. There was no evidence that this effect depends on the connectedness of a note.

One way to think of this result is that a stem and dot being connected does not serve as a basic feature forming a music note; otherwise, we should have found that music-reading expertise was correlated with the search speed difference in trials in which the music notes were all connected versus all disconnected ($r = .002$, $p = .99$, 95% CI = $[-.31, .31]$). However, it is also possible that the design in Experiment 1 did not favor the use of the connectedness information to accomplish the task, so that the participants ignored the connectedness manipulation. In Experiment 1, when the participants were looking for a valid configural target note, they were always looking for it among invalid configural distractor notes, and vice versa. In contrast, the target and distractor notes within a block were either all connected or all not connected. This manipulation might result in different sensitivities in selecting useful features from a note for the purpose of the visual search. Specifically, in Experiment 1, connectedness was not diagnostic for search.

To fully examine whether music-reading experts are sensitive to connectedness, in Experiment 2 participants were asked to detect a connected target note among disconnected distractor notes, or the other way around, while the validity of the dot–stem configuration of all notes remained consistent in a whole block. This reverse design made connectedness a diagnostic feature for the visual search, allowing us to test whether or not connectedness itself can lead to a distractor familiarity effect. Specifically, we hypothesized that, if connectedness is a defining feature of musical notation, then higher music-reading expertise would predict a more efficient search in detecting targets among distractors with valid connectedness than when looking for targets among disconnected distractors, resulting in a larger distractor familiarity effect.

Experiment 2

Method

Participants

Fifty-five people (19 males and 36 females) from the Vanderbilt community participated in the study for either

course credits or monetary reward. We were guided by the same power analysis as in Experiment 1, with support in Experiment 1 that the effect size of the correlation with expertise in musical notation was $> .4$. However, we made a greater effort to reach out to individuals with musical experience and ended up with more than 40 participants (data analysis did not proceed until the data collection had ended). Their ages ranged from 18 to 33 ($M = 21.07$ years, $SD = 3.33$). All of them had normal or corrected-to-normal vision. According to the participants' self-reports, 21 had music-reading expertise (8 males and 13 females), and 34 did not (11 males and 23 females).

Design, stimuli, and procedure

The design, stimuli, and procedure in Experiment 2 were identical to those aspects of Experiment 1, except for the following two adjustments: (1) For the music note search task, recall that in the same block of Experiment 1, the participants looked for a target note that was the only one that had a valid or invalid dot–stem configuration, while the notes on screen were either all connected or all disconnected. Therefore, regardless of what target the participants were searching for, the connectedness information was not diagnostic to identify the target, making it easier to ignore. In Experiment 2, the participants searched for a target note that was either the only note that was connected or the only note that was disconnected, while all the notes within a search display had the same validity of their dot–stem configuration. (2) For the sequential-matching task, we included shape sequences as a new control stimuli (see Fig. 2c for an example). We created ten common shapes and paired them with their mirrored form, rotated form, or a visually similar shape in order to generate the shape sequences (Fig. 7). The reference shape sequences consisted of three distinctive shapes. The test sequences were made by randomly replacing one of the three shapes with its paired shape. With this additional control condition, in Experiment 2 there were a total of six estimation series, as the exposure threshold for each stimulus was derived from the average of two estimates. Since in Experiment 1 the participants had been less variable in their performance when they matched digit sequences than when they matched note sequences (Fig. 8), this control task should capture more of the variance in performance that was not specifically tied to the musical stimuli.

Results

Perceptual fluency We first tested the correlation between the two estimates for music note sequences and the control sequences. The two estimates were positively correlated with each other for all sequences (music note sequences: $r = .50$, $p < .001$, 95% CI = $[.21, .68]$; digit sequences: $r = .69$, $p < .001$, 95% CI = $[.28, .68]$; shape sequences: $r = .27$, $p < .05$, 95% CI = $[.01, .50]$), indicating reliability of the estimation

Fig. 7 Ten pairs of common shapes used to generate the shape sequences for Experiment 2

procedure. The corrected r values based on the Spearman–Brown prophecy formula (Brown, 1910; Kaplan & Saccuzzo, 2017; Spearman, 1910) for all three correlations also confirmed the reliability of the estimation procedure (music note sequences: $r' = .67$; digit sequences: $r' = .82$; shape

sequences: $r' = .43$). The two estimates for each type of stimuli were again averaged to derive a single threshold, except for two participants whose threshold estimates were extreme relative to the others (a mean threshold of 11,831 ms for shape sequences, and a mean of 6,747 ms for note sequences, both of

Fig. 8 Box plots of the note and control thresholds, for demonstration of the distributions in Experiments 1 and 2

which fell six standard deviations away of their corresponding mean)—in those cases, the single other estimate was used. The participants had a mean note threshold of 961 ms, with a standard deviation of 472 ms; a mean digit threshold of 432 ms and a standard deviation of 262 ms; and a mean shape threshold of 842 ms and a standard deviation of 504 ms (Fig. 8). We averaged the digit threshold and the shape threshold to serve as an aggregated control threshold for each participant ($r = .36, p < .15, 95\% \text{ CI} = [.11, .57]$). Again, we confirmed that after regressing out the control-condition thresholds, the first note estimate predicted the second note estimate: $r = .59, p < .001, 95\% \text{ CI} = [.39, .74]$ (see Fig. 4). The perceptual fluency task provided a reliable measure of variance that was selectively tied to our domain of interest, music sequence matching. Finally, we calculated an expertise index for each participant that was defined as the inverse of their mean note threshold, regressing out the inverse of their mean control condition threshold (the average of the mean exposure duration thresholds for digits and shapes), in order to predict their performance in the subsequent music note search task. The inverse transformation was performed so that higher values would represent better performance.

Music note search task The accuracy in the task was high ($M = .87, SD = .11$). Following the same criterion as in Experiment 1, we used correct reaction times that were between 200 and 8,000 ms for the further analysis (a total of 13.29% incorrect trials [12.92%] and outlier data points [0.37%] were excluded from the analysis). Reaction times were log-transformed in order to correct the positive skew of the original distribution.

We first examined the relationship between one's music-reading expertise and the distractor familiarity effect, as defined by the connectedness of a note. We found no significant correlation between one's expertise index and the reaction time difference in searching for a connected target note among disconnected distractors (a familiar target among unfamiliar distractors) versus searching for a disconnected target note among connected distractors (an unfamiliar target among familiar distractors) ($r = .16, p = .24, 95\% \text{ CI} = [-.11, .41]$). As in Experiment 1, we computed the skipped correlation to ensure that potential outliers did not influence the data structure, and the results remained the same ($r_{\text{skipped}} = .11, p = .42, 95\% \text{ CI} = [-.18, .41]$). To rule out the possibility that the findings did not result from the change in the expertise index (with an additional control condition), we repeated the same analysis using the residual scores after regressing out only the variances for digit sequences, as in Experiment 1, and still obtained a null result ($r = .17, p = .21, 95\% \text{ CI} = [-.10, .42]$). No familiarity effect arising from the manipulation of the connectedness information was revealed in Experiment 2, which is consistent with the results of Experiment 1 in revealing no sensitivity to connectedness. Furthermore, we did not find any evidence for a relationship between the expertise index and the effect of dot-

stem configuration when that feature was not diagnostic ($r = .05, p = .72, 95\% \text{ CI} = [-.31, .22]$). Reaction times for detecting invalid configural targets were not predicted by the expertise index ($r = .06, p = .66, 95\% \text{ CI} = [-.32, .21]$). Note that all analyses were also done on error rates, but no significant results were found ($r_s < .09, p_s > .51$).

Discussion

In Experiment 2 we examined whether connectedness is an important basic feature in musical notation when it is a diagnostic feature for the task, and if so, whether dot-stem configuration would modulate its effect. Altogether, the results showed that perceptual fluency in reading music notes was not related to the difference in searching for a connected or disconnected target note. This suggested that whether or not the dot and stem of a note are attached does not serve as a basic feature for forming a music note, even when it was a diagnostic feature in the experimental design. Also, when dot-stem configuration was not important for identifying the target anymore, in the present design, the validity effect observed in Experiment 1 no longer existed in Experiment 2, indicating that whether a feature is diagnostic for the task demand appears to be critical in revealing distractor familiarity effect of music-reading expertise.

General discussion

In the present study we aimed to identify the defining features of music notes that characterize the boundaries of the domain of music-reading expertise. Specifically, we focused on the correct/incorrect way to form a music note, in terms of how the stem and the dot should be connected (the “dot-stem configuration” of a note) and whether or not the stem and the dot of the music note are connected (the “connectedness” of a note). In two experiments, we used a sequence-matching task in which participants matched music sequences and digit sequences (in Exps. 1 and 2), as well as shape sequences (in Exp. 2), to quantify one's expertise in reading music. With a music note search task, we asked whether we could induce a distractor familiarity effect with notes that had the valid dot-stem configuration and whose dot and stem were connected. To sum up the results, we found that expertise in reading music notes predicted the distractor familiarity effect driven by the validity of dot-stem configuration, when dot-stem configuration was a diagnostic feature for the search and connectedness remained constant (Exp. 1). Music-reading experts exhibited more efficient search for detecting an invalid configural target note among valid distractor notes, as opposed to searching in the opposite condition. This distractor familiarity effect of validity of the dot-stem configuration was robust, regardless of the connectedness information. However, experts were not sensitive to the connectedness

information when it was the relevant feature for search (Exp. 2). Results altogether suggest that dot–stem configuration but not connectedness is a basic feature in the domain of musical notation.

One difference between the present experimental setting and those in past research used for testing distractor familiarity effect is that two features of interest were tested together in our study. For example, Malinowski and Hübner (2001) asked participants to search for an *N* among mirror-*N*s, or vice versa, where the target that represented the basic feature was always relevant for the search. Critically, in each of our experiments, only one feature was tested (e.g., dot–stem configuration), while the other feature (e.g., connectedness) was also present but remained constant for every trial. It is notable that we found no evidence that the search-irrelevant feature ever influenced the efficiency of search.

One concern could be that dot–stem configuration and connectedness as presented in the present experiments might not have been equally salient in their attentional and/or perceptual properties—and we acknowledge that no effort was made to equate them. Although the reaction times for detecting a target in both experiments did not differ significantly [$t(93) = 0.32$, $p = .75$], mean accuracy was significantly higher in Experiment 1 than in Experiment 2 [$t(93) = 2.88$, $p < .01$]. It is not clear why the judgments in Experiment 1 were easier than those in Experiment 2. The connectedness manipulation might be more subtle, as it is based on only a small gap between the dot and stem, whereas dot–stem configuration can be appreciated at the scale of the entire note. The difference might also arise because configural information is fundamental and crucial for object recognition in general. Regardless of the reason, it is important to note that the two features being different in saliency does not affect our conclusions as the participants were clearly able to discriminate notes using the two features as shown in Experiments 1 (93% correct) and 2 (87% correct), and we were concerned with a potential interaction with expertise in each experiment.

The finding that dot–stem configuration elicited familiarity in music-reading experts only when it was important to the current task goal might reflect the flexibility in strategy selection for experts to use domain-specific information. This finding is in line with previous evidence that music-reading expertise may depend on the expectation to process sequences of several music-related stimuli. For instance, in Wong, Peng, Fratus, Woodman, and Gauthier (2014), music-reading experts and novices judged whether a dot presented at the fovea was on or off a line on a five-line staff that was either horizontal or vertical, where the staff orientation was manipulated in three blocking conditions, such that orientation was randomized (randomized condition), kept the same within a block (blocked condition), or alternated between horizontal and vertical (predictable condition). Experts outperformed novices only when the staff was horizontal under blocked

and predictable conditions, suggesting the flexible nature of music-reading experts' preparation for processing musical stimuli. Expertise with musical notation appears to depend on a combination of task-set and bottom-up factors.

Our results might appear inconsistent with Becker (2018), who did not find evidence that music readers and non-music-readers process musical notes differently, also using a search asymmetry design. In Becker's study, the music readers and non-music-readers looked for a regular eighth note among mirrored eighth notes. Surprisingly, both groups found the unfamiliar targets faster than the familiar targets, and the magnitude of the effects did not differ. One possibility is that by not using a quantitative measure of note-reading expertise, but instead using dichotomized groups, that study might have lacked sensitivity (MacCallum, Zhang, Preacher, & Rucker, 2002).

Indeed, one of our goals in this work was to improve on the measurement of music-reading expertise. It is not easy to derive a good definition of music-reading expertise. Since music playing is generally associated with music reading, expertise in music reading could be defined by one's music-playing background, in terms such as years of music experience and lessons (Brochard, Dufour, & Despres, 2004; Ericsson, Krampe, & Tesch-Römer, 1993; Meinz & Hambrick, 2010; Stewart, Walsh, & Frith, 2004) and/or whether one's profession is a musician or a music student (Salis, 1980). However, this sort of experience might not tell us specifically about visual abilities for processing musical notation. Following previous work (Wong & Gauthier, 2010a, 2010b, 2012; Wong et al., 2014; Wong & Wong, 2018), in the present study, music-reading expertise was defined as the perceptual fluency to read short musical sequences, and we adopted a sequence-matching task to measure participants' perceptual fluency in matching music note sequences versus one control sequence. In Experiment 1, we used digit sequences to partial out variance in performance that might not be music-related. However, perhaps because people normally have a lot of experience recognizing digits, matching digit sequences produced limited variability (see Fig. 8). Therefore, in Experiment 2, two control sequences (digit and shape sequences) were used, to capture more variance in participants' matching performance. Matching shape sequences was more comparable in its variability to note matching (see Fig. 8), but the two note thresholds still correlated highly after controlling for variance in nonnote matching, suggesting music-specific individual differences. It would have been ideal to use the improved definition of expertise in Experiment 1, and as a result, it will be important to replicate the link between expertise and sensitivity to dot–stem configuration using the new measure.

Although we have improved this measure of music-reading expertise, there are limitations to our work. First, more diverse control stimuli, other than digits and shapes, could be used in

this same task to measure perceptual fluency that is specific to musical notation domain. Second, the use of only one task to quantify an individual's ability to read music might overlook important aspects of music reading. For instance, the spatial relationship between notes is crucial for reading rhythm (the relationship between the durations of notes) and melody (the relationship between the pitches of notes). As another example, visual working memory capacity for music-related symbols could be more limited in capacity for music-reading novices than for experts. Ideally, various tasks with different task demands would be used, so that the common variance would be more likely to generalize to any task that involves musical notation reading.

To conclude, despite growing research in the domain of musical notation expertise, it is still not well established what features determine a music note for music-reading experts. In this study we found that music-reading experts are sensitive to how the dot and stem of a music note are configured. This shows that configural information is essential in efficient music reading, despite the fact that this feature is irrelevant for either rhythm or melody. Experience with nondiagnostic features can change perception, and future work should investigate how this could improve the processing of diagnostic information. Our findings also provide a new visual control appropriate for future investigations with musical notation.

Author note This work was supported by the Program for Music, Mind, and Society at Vanderbilt University and by the Government Scholarship to Study Abroad from the Ministry of Education in Taiwan. We thank Jackie Floyd and Taylor Griffith for assistance with data collection, and Yetta Wong for generously providing her stimuli. Neither of the experiments was preregistered, but the data are available online (<https://figshare.com/s/4302bec93bed149afe3e>) and the materials are available upon request.

References

- Becker, L. (2018). Familiarity affects the same event-related brain potential components in note readers and non-note readers. *Music & Science*, 1, 2059204318778237. <https://doi.org/10.1177/2059204318778237>
- Brochard, R., Dufour, A., & Despres, O. (2004). Effect of musical expertise on visuospatial abilities: Evidence from reaction times and mental imagery. *Brain and Cognition*, 54, 103–109. [https://doi.org/10.1016/S0278-2626\(03\)00264-1](https://doi.org/10.1016/S0278-2626(03)00264-1)
- Brown, W. (1910). Some experimental results in the correlation of mental abilities. *British Journal of Psychology*, 3, 296–322. <https://doi.org/10.1111/j.2044-8295.1910.tb00207.x>
- Bukach, C. M., Phillips, S. W., & Gauthier, I. (2010). Limits of generalization between categories and implications for theories of category specificity. *Attention, Perception, & Psychophysics*, 72, 1865–1874. <https://doi.org/10.3758/APP.72.7.1865>
- Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100, 363–406. <https://doi.org/10.1037/0033-295X.100.2.363>
- Fisher, R. A. (1921). On the probable error of a coefficient of correlation deduced from a small sample. *Metron*, 1, 3–32.
- Gauthier, I., Wong, A. C.-N., Hayward, W. G., & Cheung, O. S. (2006). Font tuning associated with expertise in letter perception. *Perception*, 35, 541–549. <https://doi.org/10.1068/p5313>
- Hole, G. J. (1994). Configurational factors in the perception of unfamiliar faces. *Perception*, 23, 65–74.
- Kaplan, R. M., & Saccuzzo, D. P. (2017). *Psychological testing: Principles, applications, and issues* (9th ed.). Belmont, CA: Wadsworth.
- MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, 7, 19–40. <https://doi.org/10.1037/1082-989X.7.1.19>
- Malinowski, P., & Hübner, R. (2001). The effect of familiarity on visual-search performance: Evidence for learned basic features. *Perception & Psychophysics*, 63, 458–463. <https://doi.org/10.3758/BF03194412>
- McGugin, R. W., McKeef, T. J., Tong, F., & Gauthier, I. (2011). Irrelevant objects of expertise compete with faces during visual search. *Attention, Perception, & Psychophysics*, 73, 309–317.
- McKone, E., Aimola Davies, A., Darke, H., Crookes, K., Wickramariyaratne, T., Zappia, S., & Fernando, D. (2013). Importance of the inverted control in measuring holistic face processing with the composite effect and part-whole effect. *Frontiers in Psychology*, 4, 33. <https://doi.org/10.3389/fpsyg.2013.00033>
- Meinz, E. J., & Hambrick, D. Z. (2010). Deliberate practice is necessary but not sufficient to explain individual differences in piano sight-reading skill the role of working memory capacity. *Psychological Science*, 21, 914–919.
- Moscovitch, M., Winocur, G., & Behrmann, M. (1997). What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. *Journal of Cognitive Neuroscience*, 9, 555–604.
- Mruczek, R. E. B., & Sheinberg, D. L. (2005). Distractor familiarity leads to more efficient visual search for complex stimuli. *Perception & Psychophysics*, 67, 1016–1031. <https://doi.org/10.3758/BF03193628>
- Pernet, C. R., Wilcox, R., & Rousselet, G. A. (2013). Robust correlation analyses: False positive and power validation using a new open source Matlab toolbox. *Frontiers in Psychology*, 3, 606. <https://doi.org/10.3389/fpsyg.2012.00606>
- Reicher, G. M., Snyder, C. R., & Richards, J. T. (1976). Familiarity of background characters in visual scanning. *Journal of Experimental Psychology: Human Perception and Performance*, 2, 522–530. <https://doi.org/10.1037/0096-1523.2.4.522>
- Richards, J. T., & Reicher, G. M. (1978). The effect of background familiarity in visual search: An analysis of underlying factors. *Perception & Psychophysics*, 23, 499–505.
- Richler, J. J., & Gauthier, I. (2014). A meta-analysis and review of holistic face processing. *Psychological Bulletin*, 140, 1281–1302. <https://doi.org/10.1037/a0037004>
- Richler, J. J., Gauthier, I., Wenger, M. J., & Palmeri, T. J. (2008a). Holistic processing of faces: Perceptual and decisional components. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 328–342. <https://doi.org/10.1037/0278-7393.34.2.328>
- Richler, J. J., Tanaka, J. W., Brown, D. D., & Gauthier, I. (2008b). Why does selective attention fail in face processing? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 1356–1368. <https://doi.org/10.1037/a0013080>
- Rousseeuw, P. J. (1984). Least median of squares regression. *Journal of the American Statistical Association*, 79, 871–881.
- Rousseeuw, P. J., & Van Driessen, K. (1999). A fast algorithm for the minimum covariance determinant estimator. *Technometrics*, 41, 212–223.

- Salis, D. (1980). Laterality effects with visual perception of musical chords and dot patterns. *Perception and Psychophysics*, 28, 284–292.
- Shen, J., & Reingold, E. M. (2001). Visual search asymmetry: The influence of stimulus familiarity and low-level features. *Attention, Perception, & Psychophysics*, 63, 464–475. <https://doi.org/10.3758/BF03194413>
- Spearman, C. (1910). Correlation calculated from faulty data. *British Journal of Psychology*, 3, 271–295. <https://doi.org/10.1111/j.2044-8295.1910.tb00206.x>
- Stewart, L., Walsh, V., & Frith, U. (2004). Reading music modifies spatial mapping in pianists. *Perception & Psychophysics*, 66, 183–195.
- Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory & Cognition*, 25, 583–592. <https://doi.org/10.3758/BF03211301>
- Verboten, S., & Hubert, M. (2005). LIBRA: A MATLAB library for robust analysis. *Chemometrics and Intelligent Laboratory Systems*, 75, 127–136.
- Wang, Q., Cavanagh, P., & Green, M. (1994). Familiarity and pop-out in visual search. *Perception & Psychophysics*, 56, 495–500. <https://doi.org/10.3758/BF03206946>
- Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, 5, 495–501. <https://doi.org/10.1038/nrn1411>
- Wong, Y. K., & Gauthier, I. (2010a). 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. (2010b). A multimodal neural network recruited by expertise with musical notation. *Journal of Cognitive Neuroscience*, 22, 695–713.
- Wong, Y. K., & Gauthier, I. (2012). Music-reading expertise alters visual spatial resolution for musical notation. *Psychonomic Bulletin & Review*, 19, 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, 1629–1643.
- 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, 1373–1380. <https://doi.org/10.3758/s13423-018-1483-0>
- Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, 81, 141–145. <https://doi.org/10.1037/h0027474>

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