Acta Psychologica 180 (2017) 190–199

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

The effect of beam slope on the perception of melodic contour

Warren Brodsky[a,*], Yoav Kessler[b]

[a] Music Science Lab, Department of the Arts, Ben-Gurion University of the Negev, Beer-Sheva, Israel
[b] Department of Psychology & Zlotowski Center for Neuroscience, Ben-Gurion University of the Negev, Beer-Sheva, Israel

ARTICLE INFO

Keywords:
Score-reading
Music-notation
Stroop-like task
Melodic contour
Beams

ABSTRACT

Musical score reading is a complex task, which involves attending and interpreting multiple visual constituents that are graphically congested. The present investigation examined the 'beam', which although consistently found in music notation, is typically considered as providing no more information than marking metric boundaries (i.e., chunking). However, we provide evidence here that beams enhance visual perception of contour. In Study 1, a Stroop-like paradigm was used in which participants were required to judge the direction of notes or the beam in a compound figure; the two dimensions were either congruent or incongruent. A congruency effect was observed in both tasks, confirming that both notes and beam are processed automatically during score reading. In Study 2, an additional auditory stimulus was presented. The results not only replicated the findings of Study 1, but showed that beams affect both visual and auditory perception. Finally, group differences surfaced: musicians were more affected by the direction of notes than non-musicians when attending to beams, but the effect of beams on judging note direction was comparable in both groups. The implications for understanding musical score reading – specifically issues related to melodic contour – are discussed.

1. Introduction

How difficult is it to read music notation? Music has been compared to a language – albeit non-verbal in nature – with standardized grammatical structures and syntax that allow all those who are literate to understand its meanings. Like most languages, music also has animated features to personalize self-expression. Music notation is a unique reading system employing spatial position of tones embedded into graphical representations. Almost forty years ago, Sloboda (1981/2005) concluded that music notation is the symbolic temporal structure of music. The current study considers the rapid perceptual coding processes of score readers. While the psychological effectiveness of music notation is the extent to which readers are able to retrieve information about music from the score, the compactness of the system often poses problems when more than one aspect of the same event has to be noticed causing an increase in the visual density of the information. It is often suggested (e.g., Agrillo & Piffer, 2012; Benassi-Werke, Queiroz, Araujo, Bueno, & Oliveira, 2012; Cohen, Evans, Horowitz, & Wolfe, 2011) that more efficient sight-readers are those who are particularly attuned to superordinate structures with consequential economy of coding, and that such processes occur by organizing material into higher-order interrelationships which represent certain regulations and limitations leading to cognitive expectancies. However, research demonstrating such assumptions is sporadic (with the majority of studies having been implemented by Sloboda between 1970 and 1990), and for the most part investigations comparing between musicians and non-musicians have not established high levels of ecological validity (simply because empirical tasks usually require knowledge of music and performance experience). Sloboda (1984/2005) stated that finding a way to measure music reading with absolute novices is problematic as the absence of knowledge about the names, functions, and symbols of music puts non-musicians at a disadvantage causing biases and subsequent falsification of findings. With this in mind, the current study examined the effects of an auxiliary graphic constituent found in music notation (i.e., beam) on the perception of melodic contour from a series of notes (i.e., the graphic representation of sound) and tones (i.e., the actual auditory sounds themselves). Most specifically, by manipulating *beam-slope* we examined whether the direction of the beam is processed automatically as part of score reading, and whether such processing differs between expert musicians and non-musicians. The use of beams allows for highly reliable comparisons between musicians versus non-musicians, particularly because while musicians are superior in judging pitch height (and the representation of such in musical notation), beams which connect between notes are no more than a commonplace line marking of direction and slope to which both musicians and non-musicians have an equivalent everyday knowledge and experience for decoding plane (such as ascending, descending, and maintaining a horizontal level).

* Corresponding author at: Music Science Lab, Department of the Arts, Ben-Gurion University of the Negev, POB 653, Beer-Sheva 84105, Israel.
E-mail address: wbrodsky@bgu.ac.il (W. Brodsky).

http://dx.doi.org/10.1016/j.actpsy.2017.09.013
Received 18 April 2017; Received in revised form 12 September 2017; Accepted 25 September 2017
0001-6918/

1.1. From neumes to Western musical notation

Predating the use of verbal language, behavioral repertoires involving toned utterances and expression through sound were learned by rote memorization as an oral tradition (Grout & Palisca, 1996). Early forms of notation were developed to prompt users of pitched intonation, and then later for melodic inflections. For example, signs were employed to accent the text; these developed into a set of topographical instructions known as punctuation. Such symbols, referred to as *neumes*, were used in Greek and Roman literature as well as in the Biblical *cantillation* of the ancient Hebrews. They not only served as an aid of interpretation (i.e., division of text into sentences, clauses, etc.) with which to increase reading clarity, but also instructed the reader as to when, where, and how to apply vocal inflections in order to heighten the emotional meanings of the scripture. In addition, signs were employed by cantors and monks as *cheironomic symbols* to phrase melodic shapes in a timely fashion; they made physical gestures employing hand movement patterns that developed into choir directing (i.e., conducting). Neumes were also inserted above written texts to indicate melodic movement; the relative direction and intervallic relation between notes eventually developed into music notation. About the 9th Century the practices of notating music began. Initially, space was not used to indicate duration, but rather shapes were employed; round, square, and diamond shaped note heads indicated duration of varying lengths. Over time several rhythmic conventions became the accepted practice as standard durations (multiples of each other) that setup rhythmic groupings separated by metric bar lines known as *meter*. Notation became more of a general custom about the late 12th Century with an increasing focus on polyphonic music, which rose to even greater heights in the 14th Century when the performance of notated synchronized parts became fashionable. Then, alongside the rise of instrumental music in the 16th Century, notation became more accurate with detailed specifications by composers written inside the score as guidelines to perform their works. Although some music notation systems such as Ancient Greek and Chinese remained *phonetic* (i.e., sounds represented by numbers, letters, or signs), Western music notation is *diastematic* or *intervallic* (i.e., sounds represented graphically). Readers are referred elsewhere for a musicological outline covering the development of Western music notation (see: Gorog, 2015; Strayer, 2013).

The *Orthochronic System* (OS) has been the major Western notational system for over 450 years. A central feature of OS is its abstractness, and while it does not denote a specific instrument, it does identify pitch and rhythmic relationships between notes and groups of notes. Sloboda (1981/2005) claimed this character may explain why the system has endured. OS has relevance to all musicians regardless of instrument, the historic period of music, or the music style of the repertoire performed. Nonetheless, such generality is exactly why it is necessary for each instrumentalist to have additional symbols indicating numbers for fingerings, or how to execute performance (i.e., finger stops, peddling, bowing, sticking, blowing, breathing, etc.). Moreover, a set of non-instrument-specific universal-symbols are employed for performance commands such as the nature of the attack, loudness, and phrasing. Certainly, when too many details clutter a score, readers are easily burdened. Hence, one of the strengths of OS concerns spatial constraints for simplicity of transparency. Most noticeably is the overarching employment of five horizontal lines on which note heads are placed to represent pitch height. Namely, the codification of pitch frequency is based on the spatial position of the tone on the musical stave, while a graphic set of note-head permutations designate the temporal flow of time placed within a structure that is metrically divided by bar lines.

OS is indeed a matchless system that not only integrates pitch dimensions such as frequency, duration, and volume, but also involves spatial-temporal organization as represented by single graphic symbols. Akiva-Kabiri and Henik (2012, 2014) acknowledged that the spatial positions of tones embedded in graphical representations provide a huge amount of information that is processed automatically. The first to delineate distinctions between language notation (i.e., text) versus music notation (i.e., score) was Sloboda (1981/2005) who outlined four characteristic differences: (1) whereas a text portrays a single sequence of events, a score must be able to specify different events occurring at the same time (i.e., parallel streams of information); (2) whereas text-readers are mainly concerned with understanding and remembering what they read, score-readers are essentially concerned with performing, and therefore score layout is much more important than textual arrangement or font design; (3) whereas text-readers are able to pace their own reading to accommodate the layout, score-readers cannot lose their place or experience ambiguity not even for a minute if they are to maintain the temporal flow of the performance, and hence spacing layout is far more significant for a score; and (4) whereas the position of one letter in relation to its neighboring letter is trivial in a word text, a score presents readers with complex spatial constraints at the microscopic level causing the reader to consider the positioning of each note in respect to its neighboring note.

1.2. Music reading: a task in pattern recognition

Music reading is essentially a task of recognizing familiar musical configurations as printed on the page (Waters & Underwood, 1999; Wolf, 1976). The notes readers see are essentially building blocks of larger units. Even before musicians have played a single note, they become aware of many familiar patterns, simply by searching for visual cues in the score. Musicians are so familiar with these configurations that they do not seek them consciously, but rather are processed automatically. In fact, musicians generally see a few cues and fill them with what seems appropriate to complete the pattern.

As a system, OS has to be as compact as possible, and the denseness of material poses a problem when more than one aspect of the same event has to be noticed. Therefore, a host of accepted structural conventions are employed by theoreticians, composers, performers, and publishers; these increase retrievals of information among the visual density of the score especially when distinctive visual features of the same symbol representation might indicate different aspects of the notated tones. In this connection, Sloboda (1981/2005) commented that "musicians accustomed to reading orthochronic notation at sight become very sensitive to slight changes in notational practices… Informationally and structurally [slight changes might have] absolutely no consequences at all, but [are] psychologically disruptive… the subjective impression is of something quite wrong about them" (pg. 67). Sloboda (1976a, 1976b) asserted that the perceptual difficulties of music notation regularly surface from complications in the vertical localization of individual notes whereby they may be inferred from the surrounding context. Accordingly, there is no absolute distinction between 'correct' and 'incorrect' subcomponents of music, but rather there is a continuum ranging from 'highly likely' to 'highly unlikely'; these parameters vary among players according to their degree of familiarity with the style. Although it is usually acknowledged that skilled sight-readers process notes automatically, through extensive observation Wolf (1976) concluded that in reality much guesswork is actually involved in score-reading. The first to elucidate on Wolf's finding was Sloboda (1978a/2005) who borrowed the concept of *proof-readers' error* from the literature on text reading; namely, the tendency for incorrectly spelled words to be overlooked especially when the misspelling is trivial. Sloboda claimed that music reading does not depend upon decoding of stimulus information to build up a mental representation, but rather readers use prior knowledge and expectancy to supplement and/or replace stimulus information. Hence, as musicians become more familiar and competent score-readers, they use previous knowledge to skim over the text and predict correctly what should be played. It would seem, then, that proficient score-reading is partially based on an ability to decide probable continuations within an idiom, indicating that musicians require both implicit and explicit knowledge of music theory and literature – in addition to their ability of translating such knowledge into movements that generate appropriate sounds.

Both anecdotal evidence and empirical findings demonstrate that score-readers search out familiar constellations in the printed music long before they play them, with their eyes often 2–3 measures ahead of their hands. The *eye-hand span* approach is a concept Sloboda (1974, 1978a/2005, 1978b) adapted from *eye-voice span*; he found that efficient music readers have bigger spans for reading single line melody as measured by 6–7 notes ahead of their playing. Nonetheless, contingent on the complexity of the score (e.g., the resolution density of information compacted in the notation) and the executed tempo of the performance (e.g., the temporal flow pace or cadence), readers and players might get overloaded whereby their much-mastered eye-hand span could easily disintegrate. Previously, Sloboda (1977) demonstrated that good readers treat phrases or rhythmic figures as chunks with clearly defined visual boundaries. Namely, competent score-readers are those who are particularly attuned to important superordinate structures within musical notation that link notes together into musical units with consequential economy of coding. Thus, the recognition of patterns within a musical score (aka. *transitional probabilities*) serve to offset the effects of redundancy on memory. Nonetheless, if ability to read musical notation depends on rapid short-term memorization of notes (because the eye is reading ahead of the performance), then it would be cogent to question procedural aspects of storage. In this connection, Sloboda (1980) proclaimed that musicians store material in a non-verbal non-acoustic memory, whereby visual input has direct access to a form of representation that is not aurally based – and in this context, he suggested that visual contour plays a most important aspect of input.

1.3. Contour in reading music notation

One of the building blocks in music's structural architecture is *contour*. Contour exists in language: when people speak, some of the syllables are produced at a higher pitch while others are produced at a lower pitch, and some words are verbalized loudly while others are said quietly, and some phrases are stressed more strongly while others voiced softly. Contour is an overall sound event that functions as a curve tracking the progression over time. Similarly, contour is a highly potent aspect of music perception. Many treatises on music deal with melodic contour as an essential underlying component of organization. For example, Christ, DeLone, Kliewer, Roell, and Thomson (1972) considered melody as a line that weaves through points on a musical stave, whereby one could see the wave as possessing height and depth, and both the highs and lows depict melodic motion generating the perceptive impression about melodic shape. In an early investigation exploring the effect of contour on music reading, Sloboda (1978a, 1978b) stated "since ascending and descending sequences are of central importance in much music, the ability to detect the beginning and end of such sequences may be useful to musicians" (p. 325). Sloboda outlined two types of contour: *relative contour* is perceived from information supplied by the change in direction of three adjacent notes (i.e., detecting the angle of carryover); *absolute contour* is perceived from information supplied by pairs of adjacent notes (i.e., perceiving the height of adjacent notes to the right and left). In both cases, while readers may not know the particular notes themselves, they would still be able to sense the approximate curves of the melodic line, and therefore contour as a source of general information before knowing the exact specific notes is highly potent. In his landmark study, Sloboda (1978a, 1978b) tested for differences between accomplished musicians who were also versed in notation versus musically-illiterate non-musicians, on different attributes of contour: straight line (ascending or descending), one major change in direction ("L" shape), and 2 major changes in direction ("Z" shape). The study found that the musicians were superior at reporting absolute contour, and retained more information about the relative positioning of adjacent notes. Sloboda concluded that at the 100 ms exposure, only global information is extracted such as impressions about the relative position of adjacent notes, and such information seems to be readily available to musicians presumably because they use such cues in score reading.

Yet, considering the density of information, it would be cogent to contemplate which symbols embedded in a musical score impart contour information to musicians. For example, although Sloboda exclusively considered the notes themselves, there may, in fact, be other representations which are effective vehicles in transporting information – prompting visual perceptual awareness of direction – especially as musicians read ahead of their fingers (i.e., eye-hand span). We suggest here that the 'beam' is one such possible constituent that could enhance visual perception as it marks *frontward accelerative spatial patterns*, which to some extent reflect inertia-like forward continuation *as if* a physical object were moving along a spatiotemporal trajectory (Hubbard, 2017). Albeit, to our knowledge, beaming has not attracted attention in the music cognition literature as more than a mechanism of rhythmic organization (i.e., chunking).

1.4. Beaming in musical notation

Tones less than one quarter note duration are joined by one or more thick straight lines known as *beams*. Beams connect two or more notes which are part of the same rhythmic unit within the same measure, welding them into metric configurations that allow readers to identify the middle of a measure. Beams first appeared in music notation around 1700. The standard practice of beaming notes together into rhythmic groups of the meter allowed ease and speed of reading as the employment caused scores to be perceived as less cluttered and more interpretable. Namely, the main and perhaps only *taught* function of a beam is rhythmic organization (i.e., chunking information into clearly defined units). Among the characteristic standard practices of beam usage are: direction of context (slope); connectivity to notes (heights of stems); proximity to the midline (of the stave); and graphic design (line thickness, length, and angle of slope). Since the current study manipulated beam slope as an independent variable (reflecting empirical condition), we focus here on specific graphic characteristics.

Spreadbury (2015) points out that in the early 18th Century beams were attached to stems of a standardized length measured as 3.5 stave spaces. Therefore, the visual impression was that beam lines mirrored convex, concave, or elbow shaped attachments (see Fig. 1). As can be seen in Fig. 1, the beam lines are only slightly thicker than the stem lines themselves. Another example of such practices can be seen in Fig. 2.

From the 20th century onwards a more conventional standardized use of beaming occurred. Beams became significantly thicker and bolder than stem lines. In addition, beams no longer appeared as convex, concave, or elbow shaped forms, but rather as straight lines that either sloped upwards, downwards, or remained horizontal – with all stems in the group lengthened or shortened to meet the beam above or below the rhythmic unit (see Fig. 3). We point out that the exemplar in Fig. 3 is the same as illustrated previously in Fig. 1. As can be seen in

Fig. 1. Figured Bass from opening of Sonata No. 1 Op. 2 by A. Corelli, ca. 1681, from Book of 12 Trio Sonatas, as standardly published from about 1725.

Fig. 2. Section of Prelude from Lute/Keyboard Suite in G minor BMV 995 by J. S. Bach ca. 1727–31.

Fig. 3. Figured Bass from opening of Sonata No. 1 Op. 2 by A. Corelli, ca. 1681, from Book of 12 Trio Sonatas, as standardly published in 1996.

Fig. 3, today's standard accepted practice of beaming is as follows: when the first note in the rhythmic group is lower than the last note the beam ascends (i.e., slopes upwards); when the first note in the rhythmic group is higher than the last note the beam descends (i.e., slopes downwards); and when the first and last notes match in height position the beam remains on a horizontal plane.

Considering the above, the current study explored perceptual coding processes of score readers by investigating the effectiveness of retrieving information about music from the score. Most specifically, we manipulated *beam-slope* to examine whether the direction of the beam is processed automatically as part of score reading, and whether such processing differs between readers variegated by expertise (e.g. formally-trained musicians versus musically-illiterate non-musicians). Although it could be argued that efficient sight-readers are particularly attuned to superordinate structures with consequential economy of coding, the present investigation targets the cognitive utility of beaming by hampering mental processes that are typically recruited to organize material into higher-order interrelationships. We wondered if the function of beams is solely that of structural rhythmic chunking, or if beams also serve score-readers in supplying information about melodic contour. Study 1 examined the relationship between two visual constituent dimensions found in music score – namely notes and beams. Study 2 added an aural dimension to the design that enabled us to explore cross-modality interactions between the auditory sound of the notes and the two visual constituents.

2. Study 1: unimodal visual Stroop-like task

In musical score reading, the notes always constitute the relevant dimension to which readers need direct their attention. In Study 1 we suggest that the beam direction also conveys important information about contour. However, this is done *automatically*, since attending to the beam direction is never an explicit part of note reading (for a definition of automaticity, see: Tzelgov, 1997). To demonstrate such a proposition, we employed a Stroop-like paradigm (MacLeod, 1991; Stroop, 1935). In the classic Stroop task, participants are presented with color-words, and are required to name the color ink while ignoring the word meaning. The finding is that congruency between the word and ink color (e.g., the word RED presented in font color red) leads to better performance than incongruency (e.g., the word RED presented in font color blue). This finding is typically taken as evidence for the automaticity of word reading. Stroop-like paradigms use a similar logic to examine the relationship between other dimensions, for example numerosity and size (Henik & Tzelgov, 1982). Based on current theories of expectancy regarding score reading, we created stimuli that varied in note direction (ascending pitches, descending pitches, and retaining the same pitch height), as well as beam slope (ascending line, descending line, and horizontal line). These two constituents, notes and beam, were manipulated orthogonally creating congruent (notes = beam) and incongruent (notes ≠ beam) conditions. Employing a Stroop-like task, the design enabled us to ask three questions: (a) Is the beam (direction of slope) processed automatically? If so, a congruency effect, defined as the difference in performance as measured by reaction time (*RT*) and accuracy (also known as proportion of error or *PE*) between incongruent and congruent trials, would be observed when participants judge the note direction; (b) are the notes (direction of pitches) processed automatically? If so, a congruency effect is expected when participants judge the beam direction; (c) is the automaticity of these two constituents affected by musical expertise? Since musicians are expert score readers (i.e., making judgements about note-direction), then a larger congruency effect is expected among this group when judging the direction of the beam (where the notes are irrelevant, but would nevertheless be processed presumably). Conversely, musicians and non-musicians might not differ in their expertise for judging the beam slope direction due to the fact that beams are typically thought to be irrelevant informants in score reading that solely serve the utility of marking metric groups; we note that only one rhythmic figure – four eighth-notes or quavers – was presented in the study. Accordingly, an equivalent congruency effect is expected in both groups when judging the notes (where the beam is irrelevant).

2.1. Method

2.1.1. Participants

2.1.1.1. Non-musicians. Twenty (N = 20, 17 female) undergraduate psychology majors volunteered as participants; each received extra credit points. The participants were between ages of 18–25 years old ($M = 23$, $SD = 1.60$); all reported their right hand as dominant. Half of the participants reported to have never learned an instrument, while the other half reported to have learned to play an instrument for an average of one year ($SD = 1.81$, range = 1–7). Only three participants reported to have played an instrument in the previous year, with negligible exposure to music notation ($M = 0.5$, $SD = 0.22$, range 0–1 [1 = Highly Infrequent, 4 = Highly Frequent]). Scores for basic knowledge of music notation and theory were minimal at best ($M = 2.75$, $SD = 6.73$, range = 0–20 [out of max 100 points]) as measured by the Music Notation/Theory exam (detailed below in Section 2.1.2.1).

2.1.1.2. Musicians. Twenty (N = 20, 9 female) musicians were recruited at an academy of music; they received no compensation. 70% were registered as undergraduate students in music education, performance, or theory (composition and conducting) studies. The participants were between ages of 20–47 years old ($M = 28$, $SD = 7.45$); 90% reported their right hand as dominant (but 95% preferred using their right hand during the experiment). All participants reported to have played an instrument for the last 3–32 years ($M = 13.45$, $SD = 7.77$); the principal instrument most often cited was piano (40%). All reported to have played an instrument in the previous year, with a very high frequency of exposure to music notation ($M = 3.30$, $SD = 0.86$, range 1–4), and had advanced familiarity of music notation and theory (scores in the Music Notation/Theory exam were: $M = 97.5$, $SD = 5.44$, range = 80–100).

Comparison between the two samples indicate statistically significant differences of age ($F_{(1, 38)} = 7.12$, $MSe = 29.08$, $p = 0.011$, $\eta_p^2 = 0.16$), gender proportions ($p = 0.012$), instrument performance experience ($F_{(1, 38)} = 48.64$, $MSe = 31.87$, $p < 0.001$, $\eta_p^2 = 0.56$), and scores from the music notation and theory exam ($F_{(1, 38)} = 2396.12$, $MSe = 37.46$, $p = 0.001$, $\eta_p^2 = 0.98$). The latter two differences validate the distinction between the two groups.

2.1.2. Measures, hardware, & stimuli

2.1.2.1. Music notation and theory test. Sample inclusion criteria were based on a short Music Notation/Theory exam. This evaluation was comprised of five areas, each accruing 20 points for a total max score of 100. The areas were: (1) *Notes* – identify 16 pitches on a G-clef stave by name; (2) *Chords* – identify eight tri-chords as either major or minor types on a G-clef stave; (3) *Tonality* – identify both major and minor tonalities of two key signatures on a G-clef stave; (4) *Meter* – identify the meter of a 2-measure exemplar; and (5) *Rhythm* – tap-out a 4-measure exemplar (in which a meter change occurred in measure three). In the current study, participants who received scores ≤ 20 were identified as 'non-musicians', while those who received scores ≥ 80 were identified as 'musicians'.

2.1.2.2. Equipment. The experiment employed a *ThinkPad T40* (IBM) laptop, with 14.1″ TFT (SXGA +) screen monitor, and a *SoundMax* on-board sound card. The experiment was designed, programmed, and run with *E-Prime* (Psychology Software Tools, Inc.). For the purpose of increased attentiveness, reduced environmental noise, and empirical consistency (with Experiment 2), participants wore *RH-5MA* (Yamaha) supra-aural over-the-ear semi-closed professional studio monitor headphones.

2.1.2.3. Music stimuli. Thirty 4-tone strings (hereafter referred to as *tetra-chords*) were audio recorded with an *SX-P50* (Technics 2001) 88-key touch-sensitive digital piano to a *D1600* (Korg 2001) multi-track digital recording studio desk; the exemplars were captured at a sampling size of 16-bits at a rate of 44 kHz, and saved as mono .wav sound files. The files were cropped with *Sound Forge* (Sonic Foundry 2000), and standardized for volume. The thirty tetra-chord set was comprised of ten exemplars of ascending pitches (from middle C^4-F^4 to E^5-A^5), ten exemplars of descending pitches (from F^4-C^4 to A^5-E^5), and ten exemplars of pitches retaining the same height (from D^4 to F^5); all of these represent a diapason between Middle C^4 (piano key #40, 261.626 Hz) and A^5 (piano key #61, 880 Hz). Music notation of the sound files were generated with *Finale* (MakeMusic 2005) as 24-bit .bmp picture files; each picture was a figure of four eighth-notes (quavers) coupled by a beam above or below the note head, embedded on a 5-line staff without clef, key signature, or bar line. These were presented in a standardized measure width using an enhanced engraver-quality graphic resolution (see Fig. 4). As can be seen in Fig. 4, stems are attached to the right of the note-head when beams are coupled from above, while attached to the left of the note heads when beams are coupled from below. Each of the thirty tetra-chords was coupled to three different beam genres – ascending, descending, or horizontal – producing a test-set of 90 combinations (i.e., permutations that could be seen as either congruent or incongruent).

2.1.3. Procedure

Prior to the experiment the study was approved by a review board for ethical treatment of human subjects. Participants were tested individually; they were informed that the study investigated music notation reading. Experiment 1 solely presented visual stimuli. Participants first completed a small questionnaire outlining descriptive details. They were seated approximately 60 cm in front of a visual display, and instructed to press the corresponding key as quickly as possible; they responded with the "J" key for descending notes/beam, the "K" key for notes/beam remaining the same (i.e., horizontal), and the "L" key for ascending notes/beam; they were also instructed to be careful and avoid mistakes. Each participant then completed 20 practice trials, and 180 experimental trials in each of the two attention tasks. In the Note Task they were directed to pay exclusive attention to the notes, whereas in the Beam Task they were directed to pay exclusive attention to the beams. Both condition and item order were presented at random. Each trial began with a 5-mm fixation-cross in the center of the screen for 100 ms. A music figure was then displayed in the center (32 mm [w] × 12 mm [h]); the figure remained on the screen until key-press, which initiated the subsequent figure to appear. In between the two task blocks, a second set of instructions and practice trials appeared. After every incorrect response, a visual feedback 'error' sign was provided. The experimenter remained silent throughout.

2.1.4. Analyses

Error trials were removed from the RT analysis. Also, to minimize the influence of outliers, RTs shorter than 100 ms or longer than 4000 ms were removed. This led to discarding 2.6% of the trials.

2.2. Results and discussion

2.2.1. Reaction time

An analysis of variance (ANOVA) was conducted with Group (musicians, non-musicians) as a between-subject variable, and Task (beam, notes) and Congruency (congruent, incongruent) as within-subject variables (see Table 1). The main effect of Group was significant ($F_{(1, 38)} = 16.69$, $MSe = 62{,}293.71$, $p < 0.001$, $\eta_p^2 = 0.31$), as well as the main effect of Congruency ($F_{(1, 38)} = 92.28$, $MSe = 1141.28$, $p < 0.001$, $\eta_p^2 = 0.71$). The two-way interaction between Group and Task was also significant ($F_{(1, 38)} = 8.59$, $MSe = 17{,}823.87$, $p = 0.005$, $\eta_p^2 = 0.18$). Finally, the three-way interaction was significant, $F_{(1, 38)} = 18.72$, $MSe = 888.52$, $p < 0.001$, $\eta_p^2 = 0.33$.

We continued by examining the simple interaction between Task and Congruency within each group *separately*. In non-musicians, the simple interaction between Task and Congruency was significant ($F_{(1, 19)} = 16.14$, $MSe = 877.29$, $p < 0.001$, $\eta_p^2 = 0.46$). RTs were faster in the Beam Task than in the Notes Task (788 ms versus 885 ms; $F_{(1, 19)} = 7.36$, $MSe = 25{,}630.34$, $p = 0.014$, $\eta_p^2 = 0.28$). A congruency effect was observed in both tasks (Beam: $F_{(1, 19)} = 8.25$, $MSe = 636.35$, $p = 0.010$, $\eta_p^2 = 0.30$; Notes: $F_{(1, 19)} = 34.71$, $MSe = 1670.02$, $p < 0.001$, $\eta_p^2 = 0.65$). The interaction reflected the fact that the congruency effect was larger in the Notes Task than in the Beam Task (76 ms versus 23 ms, respectively).

In the musicians group, the simple interaction between Task and Congruency was also significant ($F_{(1, 19)} = 4.47$, $MSe = 899.76$, $p = 0.048$, $\eta_p^2 = 0.19$). Unlike the non-musicians group, RTs in the Notes Task were faster than in the Beam Task, although this difference did not reach significance (662 ms versus 688 ms; $F_{(1, 19)} = 1.42$, $MSe = 10{,}017.40$, $p = 0.25$, $\eta_p^2 = 0.07$). In contrast to the non-musicians group, the congruency effect was larger in the Beam Task (67 ms; $F_{(1, 19)} = 34.54$, $MSe = 1310.14$, $p < 0.001$, $\eta_p^2 = 0.65$) than in the Notes Task (39 ms; $F_{(1, 19)} = 34.18$, $MSe = 443.10$, $p < 0.001$, $\eta_p^2 = 0.64$).

In other words, RTs in both groups were sensitive to congruency. However, non-musicians showed a greater interference from the beam when performing the notes task, while musicians were more interfered from the notes when performing the beam task.

2.2.2. Accuracy

A parallel ANOVA was conducted on the error proportion (*PE*) data. Only the main effect of Congruency reached significance ($F_{(1, 38)} = 13.27$, $MSe = 0.00029$, $p < 0.001$, $\eta_p^2 = 0.26$). *PE* was 2.5% in the incongruent condition and 1.5% in the congruent condition.

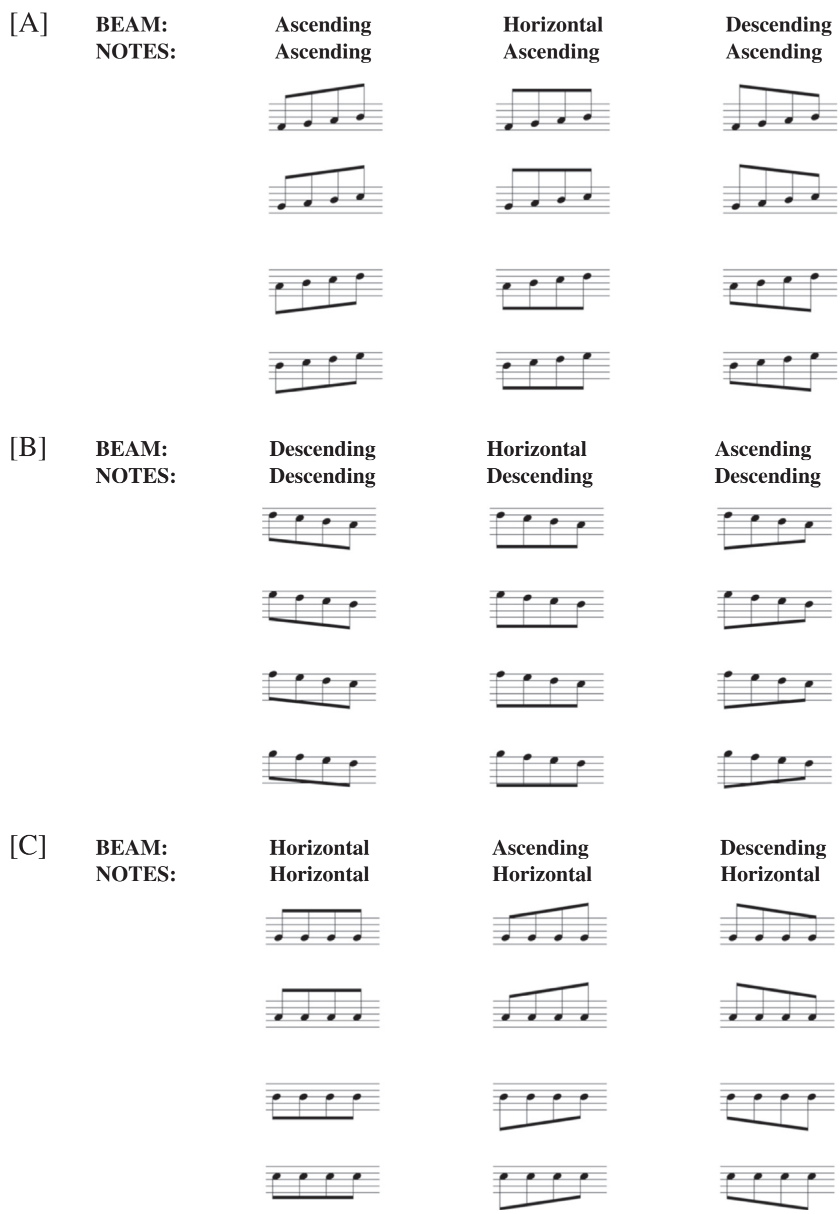

Fig. 4. The music stimuli based on combinations of beam (i.e., slope) and notes (i.e., contour).
Panel A = ascending notes; panel B = descending notes; panel C = horizontal notes.

Table 1
Study 1: Unimodal visual Stroop-like task.

Task/Congruency	Group			
	Non-musicians		Musicians	
	RT[a]	PE[a]	RT	PE
I. Beam				
Congruent[a]	776	0.02	655	0.01
Incongruent[a]	799	0.03	722	0.03
II. Notes				
Congruent	847	0.01	642	0.01
Incongruent	927	0.02	681	0.02

[a] Note: Congruent: beam = note; incongruent: beam ≠ note. RT = reaction time (ms); PE = proportion errors (*mn*).

2.3. Summary

To summarize, Study 1 demonstrated that the utility of beaming in music notation goes well beyond serving as an indicator of metrical grouping (i.e., rhythmic chunking). In our opinion, beams are graphic constituents that convey visual perceptual information cuing *frontward accelerative spatial patterns* that are required by readers of music scores. Such a concept may be linked to formulations recently published by Hubbard (2017) who claimed that music can be viewed as "tracing a trajectory through a representational space involving spatiotemporal information" (p. 21). Hubbard further pointed out that music representations reflect an inertia-like forward continuation *as if* a physical object were moving along a spatiotemporal trajectory. Accordingly, effects of momentum are experienced in both space and time. Hubbard outlined five species: Representational-momentum, Operational-momentum, Attentional-momentum, Behavioral-momentum and Psychological-momentum. Whereas the first three are primarily experienced across *space*, the latter two are primarily experienced across *time*. Regarding Musical-momentum, Hubbard contends that this species is also predominantly perceived across time, and reflects an intrinsic dynamic aspect of the mental representation that does not require a separate abstract understanding or interpretation. Considering Hubbard's proposed model, we might conclude that beams – as represented by graphic lines attached to rising and falling notes – are informants of frontward accelerative momentum that could be seen (*pun intended*) as denoting inherent spatial patterns gesturing melodic contour.

Moreover, beaming might also be a component of *eye-hand span* approach to reading music notation – albeit one that was not considered by Sloboda (1978a/2005) in his depiction of score-reading expertise. Foremost, the experiment found that when both dimensions (beam and notes) were congruent performance was faster and more accurate for both groups, while the opposite was true for both groups when dimensions were incongruent. Furthermore, group differences surfaced: interference of the beam was greater among musically-illiterate non-musicians, while expert literate musicians were not able to disengage from the notes (i.e., higher degree of interference for the Note Task than the Beam Task). This latter find suggests the possibility of a hierarchical taxonomy of informational processing: visual perceptive information prompted during score reading from the notes (i.e., pitch height) is far more salient than from the beams (slope direction of contour). Alternatively, the latter finding may reflect explicit learning (i.e., over-learning) of seeking information about pitch height (notes) while informational cues concerning contour (beam slope) falls within the realms of implicit learning due to practice and performance experience.

3. Study 2: cross-modal visual-aural Stroop-like task

When considering that music notation is the visual graphic representation of aural impressions subsequent to perception of an acoustic event, Experiment 2 added aural stimuli to the empirical platform. We expected that a cross-modal visual-aural Stroop-like task would increase attentional effects (facilitation of responses when note/beam/sound is congruent, and inhibition of responses when note/beam/sound are incongruent). Finally, as one level of music reading can be seen as an act of deciphering notation, and a second level is the conversion of visual images to acoustic events, the current study explored five different combinations of the compound stimuli. Subtype 1, *Full Congruency*: beam = note = sound; Subtype 2, *Partial Congruency*: (beam = note) ≠ sound; Subtype 3, *Partial Congruency*: (beam = -sound) ≠ note; Subtype 4, *Partial Congruency*: (note = sound) ≠ beam; and Subtype 5, *Full Incongruency*: beam ≠ note ≠ sound. It should be pointed out that commonalities among Subtypes 1 & 3 & 4 reflect cross-modal congruencies between visual-aural dimensions, while Subtypes 2 & 5 reflect cross-modal incongruencies.

3.1. Method

3.1.1. Participants

3.1.1.1. Non-musicians. Twenty (N = 20, 15 female) undergraduate psychology majors volunteered as participants; each received extra credit points. The participants were between ages of 21–28 years old ($M = 24$, $SD = 1.58$); 85% reported their right hand as dominant. 60% of the participants reported to have previously learned an instrument; the duration of formal learning was on average for a half year ($SD = 0.55$, range = 0–2). Only three of these participants reported to have played an instrument in the previous year, with negligible exposure to music notation ($M = 0.25$, $SD = 0.55$, range 0–2 [1 = Highly Infrequent, 4 = Highly Frequent]). Scores for basic knowledge of music notation and theory were minimal ($M = 4.19$, $SD = 9.99$, range = 0–37.5 [out of max 100 points]).

3.1.1.2. Musicians. Twenty (N = 20, 8 female) musicians were recruited from a music academy; they received no compensation. 95% were undergraduate students in music education, performance, or theory (composition and conducting) studies. The participants were between ages of 18–30 years old ($M = 23.3$, $SD = 2.98$); 75% reported their right hand as the dominant hand (but all 100% preferred to use the right hand during the experiment). All participants reported to have played an instrument for the last 7–22 years ($M = 14.35$, $SD = 4.18$); the principal instrument most often cited was piano (50%). All reported to have played an instrument in the previous year, with a very high frequency of exposure to music notation ($M = 3.75$, $SD = 0.79$, range 1–4), and demonstrated an advanced familiarity of music theory ($M = 96.88$, $SD = 6.33$, range = 80–100).

Comparisons between the two samples indicate statistically significant differences of gender proportions ($p = 0.031$), instrument performance experience ($F_{(1,\ 38)} = 216.25$, $MSe = 8.90$, $p < 0.001$, $\eta_p^2 = 0.85$), and music notation exam score ($F_{(1,\ 38)} = 1229.34$, $MSe = 69.88$, $p < 0.001$, $\eta_p^2 = 0.97$). The latter two differences validate the distinction between the two groups.

3.1.2. Measures, hardware, & stimuli

The methods and equipment used in Study 2 were identical to Study 1 with one exception: the 30 4-tone strings (*tetra-chords*) presented previously as visual stimuli were also presented aurally as sound clips to the participants. The sound duration of each audio exemplar was approximately 1 s (1000 ms).

3.1.3. Procedure

The procedures employed in Study 2 were identical to those of Study 1 with two major exceptions: (1) an aural stimulus was presented simultaneously with the visual display in each trial; and (2) we added a block in which participants had to judge the contour of the pitches (ascending, descending, or retaining the same height) presented aurally. Each Stroop-like task (beam, notes, sound) started with 20 practice trials, and was followed by 162 experimental trials. The stimuli were composed of all combinations of note direction (ascending notes, descending notes, notes maintaining a horizontal plane), beam direction (ascending slope, descending slope, slope maintaining horizontal plane), and sound direction (ascending sounds, descending sounds, sounds maintaining horizontal plane). Six stimuli were presented for each of the above 27 combinations, resulting in 162 experimental trials. The stimuli within each combination differed in the exact pitch and position on the stave. The order of the tasks was counterbalanced between participants, and order of the items was presented at random. During the task in which participants were directed to pay exclusive attention to the sound, the experimenter carefully monitored the participant's eye-gaze towards the computer screen (in order to assure that eyes were not shut).

3.1.4. Analyses

RT exclusion criteria were the same as in Study 1. This led to discarding 4.7% of the trials.

3.2. Results and discussion

3.2.1. Reaction time

An ANOVA was conducted with Group (musicians, non-musicians) as a between-subject variable and Task (Beam, Notes, Sound) and Subtype (1–5) as within-subject variables (see Table 2). All three main effects were significant (Group: $F_{(1,\ 38)} = 10.97$, $MSe = 397{,}862.71$, $p = 0.002$, $\eta_p^2 = 0.22$; Task: $F_{(2,\ 76)} = 119.68$, $MSe = 70{,}530.29$, $p < 0.001$, $\eta_p^2 = 0.76$; and Subtype: $F_{(4,\ 152)} = 27.05$, $MSe = 3285.70$, $p < 0.001$, $\eta_p^2 = 0.42$). The two-way interaction between Group and Task was significant ($F_{(2,\ 76)} = 8.98$, $MSe = 70{,}530.29$, $p < 0.001$, $\eta_p^2 = 0.19$), as well as the two-way interaction between Task and Subtype ($F_{(8,\ 304)} = 5.25$, $MSe = 3091.70$, $p < 0.001$, $\eta_p^2 = 0.12$). Finally, the three-way interaction was significant ($F_{(8,\ 304)} = 4.60$, $MSe = 3091.70$, $p < 0.001$, $\eta_p^2 = 0.11$).

We continued by examining each task separately. In the Beam Task, the simple interaction between Group and Subtype was significant ($F_{(4,\ 152)} = 11.24$, $MSe = 2463.00$, $p < 0.001$, $\eta_p^2 = 0.23$). RT did not differ between the subtype conditions for non-musicians ($F_{(4,\ 76)} = 0.96$, $MSe = 2543.85$, $p = 0.43$, $\eta_p^2 = 0.05$). However, in the musicians group significant differences were observed among the

Table 2
Study 2: Cross-modal visual-aural Stroop-like task.

Task/Subtype	Group			
	Non-musicians		Musicians	
	RT*	PE*	RT	PE
I. Beam				
Subtype 1: beam = note = sound	823	0.03	708	0.03
Subtype 2: (beam = note) ≠ sound	810	0.03	724	0.02
Subtype 3: (beam = sound) ≠ note	823	0.04	790	0.03
Subtype 4: (note = sound) ≠ beam	815	0.03	829	0.03
Subtype 5: beam ≠ note ≠ sound	828	0.03	828	0.03
II. Notes				
Subtype 1	868	0.03	687	0.01
Subtype 2	889	0.04	706	0.01
Subtype 3	945	0.06	749	0.02
Subtype 4	919	0.05	721	0.02
Subtype 5	941	0.06	761	0.03
III. Sound				
Subtype 1	1242	0.10	986	0.00
Subtype 2	1291	0.14	1063	0.02
Subtype 3	1280	0.12	1066	0.01
Subtype 4	1248	0.08	991	0.01
Subtype 5	1300	0.15	1045	0.02

Note: RT = reaction time (ms); PE = proportion of errors (*mn*).
Key: Subtype 1, *Full Congruency*: beam = note = sound; Subtype 2, *Partial Congruency*: (beam = note) ≠ sound; Subtype 3, *Partial Congruency*: (beam = sound) ≠ note; Subtype 4, *Partial Congruency*: (note = sound) ≠ beam; Subtype 5, *Full Incongruency*: beam ≠ note ≠ sound.

subtypes ($F_{(4,\ 76)} = 26.54$, $MSe = 2382.14$, $p < 0.001$, $\eta_p^2 = 0.58$). Post-hoc Tukey tests revealed that for musicians Subtypes 1 & 2 were significantly faster than Subtypes 3 & 4 & 5 (all *p*-values < 0.001). In other words, the musicians judged the beam direction faster if it was congruent with the note direction than when the beam and the notes were incongruent – regardless of the direction or compatibility with the sound. No other significant differences were observed among the subtypes. Moreover, the simple effect of Group in the Beam Task was non-significant ($F_{(1,\ 38)} = 0.77$, $MSe = 148{,}794.83$, $p = 0.39$, $\eta_p^2 = 0.02$) indicating that RT was equivalent for both groups in this task. In other words, musicians did not demonstrate an overall advantage over non-musicians when judging the beam direction.

In the Note Task, the simple interaction between Group and Subtype was non-significant ($F_{(4,\ 152)} = 0.38$, $MSe = 2966.87$, $p = 0.83$, $\eta_p^2 = 0.01$), indicating that the effects of subtype were equivalent in the two groups. The simple effect of Group was significant ($F_{(1,\ 38)} = 9.93$, $MSe = 191{,}190.97$, $p = 0.003$, $\eta_p^2 = 0.21$), demonstrating a faster performance in the musicians group (725 ms) compared to the non-musicians (912 ms). This finding could be expected due to the musicians expertise in reading music notation. Further, the simple effect of Subtype was significant ($F_{(4,\ 152)} = 14.65$, $MSe = 2966.87$, $p < 0.001$, $\eta_p^2 = 0.28$). Post-hoc Tukey tests revealed significant differences between Subtype 1 versus Subtypes 3 & 4 & 5 (all *p*-values < 0.003), and between Subtype 2 versus Subtypes 3 & 5 (both *p*-values < 0.001). This pattern demonstrates that judging the notes direction was faster when it was congruent with the beam direction, for both groups.

Finally, in the Sound Task, the simple effect of Group was significant ($F_{(1,\ 38)} = 18.19$, $MSe = 198{,}937.49$, $p < 0.001$, $\eta_p^2 = 0.32$) with quicker RTs for the musicians compared to the non-musicians (1030 ms versus 1272 ms), again demonstrating their superior performance in sound processing. The simple effect of Subtype was significant ($F_{(4,\ 152)} = 9.87$, $MSe = 4039.24$, $p < 0.001$, $\eta_p^2 = 0.21$). The simple interaction between Group and Subtype was non-significant ($F_{(4,\ 152)} = 1.33$, $MSe = 4039.24$, $p = 0.26$, $\eta_p^2 = 0.03$), indicating that the difference between the subtypes were equivalent for both groups. Post-hoc Tukey tests showed significant differences between Subtype 1 versus Subtypes 2 & 3 & 5 (all *p*-values < 0.001), and between Subtype 2 versus Subtype 4 ($p = 0.003$). All these differences demonstrate faster judgement of the sound when it was congruent with the notes. Subtype 3 was faster than Subtype 4 ($p = 0.003$), again indicating that congruence between the sound and the notes is more important for sound judgments than congruence between the sound and the beam. Finally, the difference between Subtype 4 versus Subtype 5 was also significant ($p < 0.001$).

3.2.2. Accuracy

A parallel ANOVA was conducted on the *PE* data. All three main effects were significant (Group: $F_{(1,\ 38)} = 20.87$, $MSe = 0.0179$, $p < 0.001$, $\eta_p^2 = 0.35$; Task: $F_{(2,\ 76)} = 6.82$, $MSe = 0.0104$, $p = 0.002$, $\eta_p^2 = 0.15$; and Subtype: $F_{(4,\ 152)} = 4.65$, $MSe = 0.0014$, $p = 0.001$, $\eta_p^2 = 0.11$). A significant two-way interaction was observed between Group and Task ($F_{(2,\ 76)} = 12.96$, $MSe = 0.0104$, $p < 0.001$, $\eta_p^2 = 0.25$). This interaction indicated more accurate performance in the musicians group when judging the Notes ($F_{(1,\ 38)} = 12.79$, $MSe = 0.0039$, $p = 0.001$, $\eta_p^2 = 0.25$), and the Sound ($F_{(1,\ 38)} = 17.62$, $MSe = 0.0329$, $p < 0.001$, $\eta_p^2 = 0.32$), but not when judging the Beam ($F_{(1,\ 38)} = 2.93$, $MSe = 0.0019$, $p = 0.09$, $\eta_p^2 = 0.07$). Again, this finding reflects the fact that the musicians expertise is only evident when judging the notes and sound dimensions. The interaction between Task and Subtype was also significant ($F_{(8,\ 304)} = 3.21$, $MSe = 0.0012$, $p = 0.002$, $\eta_p^2 = 0.08$). Specifically, the simple effect of Subtype was non-significant in the Beam Task ($F_{(4,\ 156)} = 0.60$, $MSe = 0.0011$, $p = 0.66$, $\eta_p^2 = 0.02$), but was significant in both Note Task ($F_{(4,\ 156)} = 3.30$, $MSe = 0.0010$, $p = 0.01$, $\eta_p^2 = 0.08$) and Sound Task ($F_{(4,\ 156)} = 5.97$, $MSe = 0.0017$ $p < 0.001$, $\eta_p^2 = 0.13$). Post-hoc Tukey tests indicated that in the Note Task, the effect stemmed from a significant difference between Subtype 1 versus Subtype 5 ($p = 0.03$), namely full crossmodal congruency and full incongruency, while in the Sound Task significant differences were observed between Subtype 1 versus Subtype 2 ($p = 0.04$), Subtype 1 versus 5 ($p = 0.004$), Subtype 2 versus Subtype 4 ($p = 0.009$), and Subtype 3 versus Subtype 5 ($p < 0.001$). The three-way interaction was non-significant ($F_{(8,\ 304)} = 1.09$, $MSe = 0.0012$, $p = 0.37$, $\eta_p^2 = 0.03$).

3.3. Summary

To summarize, Experiment 2 implemented a cross-modal Stroop-like paradigm demonstrating differences between musically-literate expert musicians and musically-illiterate non-musicians in an ecologically valid empirical task without causing disadvantage or bias towards those without expertise. The experiment confirms that music notation is, in fact, the symbolic temporal structure of music, but ultimately such perception is influenced by the visual mass of symbols employed in notation. Moreover, the findings demonstrate a set of overtly honed skills developed by musicians towards aural selective attention needed for an increased aptitude of masking streams of auditory information outside of their performance focus. Namely, they were much less hampered when judging the visual dimensions by dissimilar and contrasting parallel sound (i.e., aural streams of information) than by either dissimilar and contrasting parallel notes or beams (i.e., visual streams of information).

4. General discussion and conclusions

A music score's primary function is to assist a musician in performance (Sloboda, 1981/2005). Accordingly, performers must be able to keep their place, even if they do not use all of the information; performers must be able to find the appropriate information rapidly and effectively. Hence, important features of music notation are compactness, discriminability, and consistency. Yet, the congestion of information in a music score does not always allow such functional transparency. Therefore, in their effort to command the most substantial level of proprioceptive engagement, performers must be able to capture information by marginal cues – some of which are often thought to be inconsequential. We contend that because music performers' eyes read ahead of what their hands are playing (i.e., the *eye-hand* *span*), and since auditory perception via their ears has for all other purposes been taken hostage by acoustic informants in the environment (i.e., the sound sources of their own instrument and instruments of other performers as they monitor for synchronicity of performance), and since such a process subsequently masks the internal imagery of music generated by music notation (referred to as *notational audiation* [see: Brodsky & Henik, 1997; Brodsky, Henik, Rubinstein, & Zorman, 1998, 1999, 2003; Brodsky, Kessler, Rubinstein, Ginsborg, & Henik, 2008]), then music performers are left to seek information by way of *snapshot peeks* and *flash glimpses* at visual pointers and indicators found in the score. The current study proposes that musicians do in fact rely more on visual cues of contour to enhance their performance than has been acknowledged in the past, and to this end, graphic representations of contour may be their perceptual anchor. Beams appear above or below the actual notes, and therefore can be seen as different but parallel visual streams, and perhaps serve as highly important markers of spatial information. Nonetheless, beaming has never attracted the attention of scientists who have explored processes of reading music notation, nor have those investigating the *musical mind* found beams to be as significant as other signs or symbols in music scores. Furthermore, music theory teachers usually do not attach value to beaming beyond what is customarily seen as its sole utility – a marker of metric grouping (i.e., rhythmic chunking).

In discussing music-reading, Gudmundsdottir (2010) contends that traditional methods for teaching music-reading skills have always been flawed. Hence, fluent music literacy is usually acquired through a varied and repeated *trial-and-error* attempt at problem solving towards achieving mastery. Traditionally, the approaches employed in music-reading instruction have mostly been based on an instructor's own personal experiences, often fortified by anecdotal testimonials conveyed by well-known performers. Yet, as Gudmundsdottir brings out, problems in music-reading acquisition are more common than one may suspect, and difficulties among large numbers of music students have prompted many music educators and instrumental tutors to simply abandon music-reading instruction. Accordingly, when students fail to develop acceptable fluency in music reading, instrument teachers have little more than their own intuition to formulate a recommended strategy. We feel that such a situation might explain why a more in-depth understanding of beams is lacking, and why beams as a representation of melodic contour are so widely overlooked.

The term *music reading* implies the act of decoding the symbols of music notation using a musical instrument; it has also been referred to as *sight-reading*. Long ago, Wolf (1976) claimed that music reading is a complex process, involving at least two distinct independent skills: the reading skill and the mechanical skill. However, from a cognitive perspective, music reading requires several simultaneous processes, including: coding of visual information, motor responses, visual-motor integration, and aural perception. Moreover, there is empirical evidence by Kopiez, Weihs, Ligges, and Lee (2006) demonstrating that music-reading achievement at a high level is determined not only by psychomotor speed (i.e., ability to play the notes correctly demonstrated by performance analyses), but also by the speed of information processing (demonstrated by cognitive-based neuroscience analyses). Although motor response and decoding abilities are truly important components of music-reading, it is the integration of both facilities that is key to successful execution and mastery.

Studies on reading music notation indicate that one of the differences between experts and less proficient readers, is that the experts look further ahead in the music due to their ability to perceive the musical notation in larger chunks (Sloboda, 1974, 1978a/2005, 1978b). Chunking of information in a music score depends on the perception of identifiable clusters or entities (e.g., tonal patterns or rhythmic patterns). Success in music reading, then, seems to depend on an awareness of musical structures and the efficient capture of such information in real time – especially pertinent is temporal spatial perceptiveness concerning frontward accelerative spatial patterns (that unfold as melodic contour).

As can be expected, the current study did find that musicians were more expert in dealing with music notation than non-musicians. However, the efficiency of our empirical task in decreasing biases against the non-musicians was highly potent, and hence the investigation was implemented in an ecologically valid fashion. For example, the findings of Study 1 illustrate that overall musicians were just 0.5% more accurate and 162 ms faster than non-musicians; in the Beam Task musicians were 1% more accurate and 99 ms faster; in the Note task both groups were just as accurate albeit musicians were 223 ms faster. Yet, while RTs in both groups were seen to be sensitive to congruency, musicians were indeed more interfered by the notes when requested to allocate attention to beams, while non-musicians showed a greater interference from the beams when focusing on the notes. Perhaps, for musicians who spend their lives decoding pitch heights, beams provide less information than notes, while among non-musicians the tetrachord is no more than a figure constructed of two parallel streams of information (one represented by four dots •••• and the other by a line ——). Nonetheless, the picture became all the clearer in Study 2. When considering that music notation is the visual graphic representation of aural impressions subsequent to perception of an acoustic event, we expected that a cross-modal visual-aural Stroop-like task would increase attentional efforts and demonstrate facilitation of responses when constituents would be congruent, as well as demonstrate inhibition of responses when constituents would be incongruent. Hence, Study 2 added parallel aural stimuli to the previously employed uni-modal visual tetrachord symbol, generating a more compound and complex cross-modal visual-aural array. By adding sound to the tetrachord symbol, we were able to supplement and crowd far more information in to the exemplar. In fact, five permutations or combinations (referred to above as ‘subtypes’) surfaced between the two visual constituents and sound constituent: one fully congruent figure, one fully incongruent figure, and three figures that consisted of partial congruencies between visual and sound constituents. In general, Study 2 illustrates that responses of musicians were overall more accurate by 5% and faster by 158 ms. Musicians are indeed more expert, and such expertise widened with the degree of proficiency required by the task: musician responses were 1% more accurate and 45 ms faster in the Beam Task, 3% more accurate and 188 ms faster in the Note Task, and 11% more accurate and 242 ms faster in the Sound Task. Furthermore, there were significant variances of musician responses depending on the subtype; musicians were significantly better at deciphering figures consisting of congruencies between the visual constituents (i.e., beam = notes) rather than when one visual constituent was congruent with the sound while the other was dissimilar and contrasting (i.e., [beam = sound] ≠ note, or [note = sound] ≠ beam). Based on Gregoire, Perruchet, and Poulin-Charronnat (2014a, 2014b) we might see the underlying concept as one where greater interference occurs when responding to the target implies ‘translation’ from one code to another – for example, having to translate beams or notes or sound to a verbal code: ‘upwards’, ‘downwards’, or ‘stays-the-same’. On the other hand, non-musicians demonstrated no significant variances of responses between the subtypes. Hence, we would view these latter findings as evidence that beams do provide ample information beyond rhythmic organization (i.e., marking metric figures), and that such visual perceptive information as provided by beams related to contour – which cannot be as easily ignored as can a parallel aural stream of notes.

Acknowledgement

The authors would like to acknowledge Ms. Gali Naim (Department of Psychology, Ben-Gurion University of the Negev) for her assistance in data collection among the non-musician samples in both studies. In addition, the authors thank Professor Veronika Cohen and Dr. Tatiana Vendrova (Department of Music Education, Jerusalem Academy of Music and Dance) for assistance in recruiting musicians for both studies.

References

Agrillo, C., & Piffer, L. (2012). Musicians outperform nonmusicians in magnitude estimation: Evidence of a common processing mechanism for time, space, and numbers. *The Quarterly Journal of Experimental Psychology, 65*(12), 2321–2332.

Akiva-Kabiri, L., & Henik, A. (2012). A unique asymmetrical Stroop effect in absolute pitch possessors. *Experimental Psychology, 59*(5), 272–278.

Akiva-Kabiri, L., & Henik, A. (2014). Additional Insights. Commentary on "The musical Stroop effect: Opening a new avenue to research on automatisms" by L. Gregoire, P. Perruchet, and B. Poulin-Charronnat (Experimental Psychology, 2013, Vol 60, pp. 269–278). *Experimental Psychology, 61*(1), 75–77.

Benassi-Werke, M., Queiroz, M., Araujo, R. S., Bueno, O. F. A., & Oliveira, M. G. M. (2012). Musicians; working memory for tones, words, and pseudowords. *The Quarterly Journal of Experimental Psychology, 65*(6), 1161–1171.

Brodsky, W., & Henik, A. (1997, 7–12 June). *Demonstrating inner hearing among musicians. Paper presented at the Third Triennial ESCOM Conference, Uppsala, Sweden*.

Brodsky, W., Henik, A., Rubinstein, B. S., & Zorman, M. (1999). Inner hearing among symphony orchestra musicians: Intersectional differences of string players versus wind players. In S. W. Yi (Ed.). *Music, mind, and science* (pp. 370–392). Seoul: Seoul National University Press.

Brodsky, W., Henik, A., Rubinstein, B. S., & Zorman, M. (2003). Auditory imagery from musical notation in expert musicians. *Perception & Psychophysics, 65*(4), 602–612.

Brodsky, W., Henik, A., Rubinstein, B.-S., & Zorman, M. (1998, August 26–30). *Demonstrating inner hearing among high-trained expert musicians. Paper presented at the 5th international conference on music perception and cognition, College of Music, Seoul National University, Seoul, Korea*.

Brodsky, W., Kessler, Y., Rubinstein, B.-S., Ginsborg, J., & Henik, A. (2008). The mental representation of music notation: Notational audiation. *Journal of Experimental Psychology. Human Perception and Performance, 34*(2), 427–445.

Christ, W., DeLone, R., Kliewer, V., Roell, L., & Thomson, W. (1972). *Materials and structure of music. Vol. 1*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Cohen, M. A., Evans, K. K., Horowitz, T. S., & Wolfe, J. M. (2011). Auditory and visual memory in musicians and nonmusicians. *Psychonomic Bulletin & Review, 18*, 586–591.

Gorog, C. S. (2015). Slashes, dashes, points, and squares: The development of musical notation. Paper presented at the Research and Scholarship Symposium, Cedarville University, April 1, 2015. Retrieved from http://digitalcommons.cederville.edu/research_scholarship_symposium/2015/podium_Presentation/5.

Gregoire, L., Perruchet, P., & Poulin-Charronnat, B. (2014a). About the unidirectionality of interference: Insight from the musical Stroop effect. *The Quarterly Journal of Experimental Psychology, 67*(11), 2071–2089.

Gregoire, L., Perruchet, P., & Poulin-Charronnat, B. (2014b). Is the musical Stroop effect able to keep its promise? A reply to Akiva-Kabiri and Henik (2014), Gast (2014), Moeller and Frings (2014), and Zakay (2014). *Experimental Psychology, 61*(1), 80–83.

Grout, D. J., & Palisca, C. V. (1996). *A history of Western music* (5th edition). New York, NY: W.W. Norton & Company, Inc.

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

Henik, A., & Tzelgov, J. (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. *Memory & Cognition, 10*, 389–395.

Hubbard, T. L. (2017). Momentum in music: Musical succession as physical motion. *Psychomusicology: Music, Mind, & Brain, 27*, 14–20.

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

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*, 163–203.

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

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

Sloboda, J. A. (1976b). The effect of item position on the likelihood of identification by inference in prose and music reading. *Canadian Journal of Psychology, 30*(4), 228–237.

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

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

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

Sloboda, J. A. (1980). The psychological reality of musical segments. *Canadian Journal of Psychology, 34*(3), 274–280.

Sloboda, J. A. (1981). The uses of space in music notation. *Visible Language, 15*(1), 86–110.

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

Sloboda, J. A. (Ed.). (2005). *Exploring the musical mind*. Oxford: Oxford University Press.

Spreadbury, D. (2015). Development diary, part 10: Making notes. Retrieved from http://blog.steinberg.net/2015/03/development-diary-part-10/.

Strayer, H. (2013). From neumes to notes: The evolution of music notation. *Musical Offerings, 4*(1), 1–14. Retrieved from http://digitalcommons.cederville.edu/musicalofferings/vol4/iss1/1.

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

Tzelgov, J. (1997). Specifying the relations between automaticity and consciousness: A theoretical note. *Consciousness and Cognition, 6*, 441–451.

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.

Wolf, T. (1976). A cognitive model of sight-reading. *Journal of Psycholinguistics, 5*(2), 143–171.