BRILL *Timing & Time Perception* (*2015*) DOI:10.1163/22134468-03002039 TIME brill.com/time

The Effect of Visual Information on Young Children's Perceptual Sensitivity to Musical Beat Alignment

Kathleen M. Einarson[1] **and Laurel J. Trainor**[1, 2, 3, *]

[1] Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, ON L8S 4K1, Canada

[2] McMaster Institute for Music and the Mind, McMaster University, Hamilton, ON L8S 4K1, Canada

[3] Rotman Research Institute, Baycrest Centre, Toronto, ON M6A 2E1, Canada

Received 1 April 2014; accepted 28 October 2014

Abstract

Recent work examined five-year-old children's perceptual sensitivity to musical beat alignment. In this work, children watched pairs of videos of puppets drumming to music with simple or complex metre, where one puppet's drumming sounds (and movements) were synchronized with the beat of the music and the other drummed with incorrect tempo or phase. The videos were used to maintain children's interest in the task. Five-year-olds were better able to detect beat misalignments in simple than complex metre music. However, adults can perform poorly when attempting to detect misalignment of sound and movement in audiovisual tasks, so it is possible that the moving stimuli actually hindered children's performance. Here we compared children's sensitivity to beat misalignment in conditions with dynamic visual movement versus still (static) visual images. Eighty-four five-year-old children performed either the same task as described above or a task that employed identical auditory stimuli accompanied by a motionless picture of the puppet with the drum. There was a significant main effect of metre type, replicating the finding that five-year-olds are better able to detect beat misalignment in simple metre music. There was no main effect of visual condition. These results suggest that, given identical auditory information, children's ability to judge beat misalignment in this task is not affected by the presence or absence of dynamic visual stimuli. We conclude that at five years of age, children can tell if drumming is aligned to the musical beat when the music has simple metric structure.

Keywords

Music, beat alignment, metre, enculturation, development, multisensory perception

* To whom correspondence should be addressed. E-mail : LJT@mcmaster.ca

1. Introduction

Music is socially pervasive, and much musical experience is amassed via exposure, such as in the context of formal events like weddings, or of everyday experiences like watching television. These types of experiences do not typically involve music as an isolated auditory stimulus, but are instead multimodal experiences. For example, dancing to music requires coordination of motor movements on the part of the listener, and in live or video recorded music performances people see as well as hear performers. Despite such everyday multisensory musical experiences, people are quite tolerant of timing misalignments between auditory and visual stimuli (e.g., Phillips-Silver et al., 2011; Sekuler et al., 1997; Vatakis & Spence, 2006a, b). In a previous study (Einarson & Trainor, subm.), we found that five-year-old children were able to detect phase and tempo errors in puppets' drumming to musical excerpts with simple metrical structures. Here we replicate this finding and investigate whether concurrent dynamic visual information affects children's performance on this task.

Musical timing has a *grouping* structure, which corresponds to the perception of phrases, as well as a *metrical* structure, which corresponds to the regular underlying beats and can be perceptually abstracted from the rhythmic surface (Gjerdingen, 1989; Lerdahl & Jackendoff, 1983). Rhythm and metre are universal aspects of music that vary between and across human cultures, and this paper is concerned with perception of metrical structure. Metrical structure is hierarchically organized, with the beats at different levels occurring at faster or slower rates (Essens & Povel, 1985). Some beats are accented and thus are perceived as stronger than others at a given level (Jones, 1976) and only accented beats are passed to the next level of the metrical hierarchy (Essens, 1986).

In simple metrical structures, accented beats occur at regular time intervals at each level of the hierarchy. Such regularly spaced beats are referred to as isochronous. Metre in Western music typically has simple metric structure, and most commonly possesses an underlying binary structure with a 2:1 ratio between beats at successive hierarchical levels (London, 2004). When notated, this music has time signatures such as 4/4 or 2/4. However, not all music is composed of a pattern that produces equally spaced pulses at each level of the hierarchy. For example, at one level of the metrical hierarchy, accents can occur to form patterns containing successive groups of two or three beats and, thus, irregularly spaced accents. These complex metrical structures are referred to as non-isochronous. Music notation reflects these complex patterns using notation such as 5/4 and 7/8.

Although Western music is primarily composed in simple metres, music from many other cultures (e.g., South Asia, Africa, the Middle East, and the Balkan region of southeastern Europe) commonly employs complex metrical structures. Adults from cultures such as Bulgaria and Macedonia demonstrate equal perceptual sensitivity to both simple and complex metrical stimuli

(Snyder et al., 2006), and in cultures where complex metric structures are common, children also learn them from a young age, seemingly without difficulty (Rice, 1994). Infants are also sensitive to the beat of music very early in development (Patel et al., 2011; Winkler et al., 2009; Zentner & Eerola, 2010), and exposure over the course of the first year after birth to culturally typical metrical structures results in perceptual specialization for those metres (Hannon & Trehub, 2005; Soley & Hannon, 2010). This experience-guided perceptual narrowing in infancy is analogous to that seen across many domains, including the perception of phonemes (e.g. Trehub, 1976), faces (Kelly et al., 2007), voices (Friendly et al., 2013, 2014), and musical tonal structures (Hannon & Trainor, 2007; Lynch & Eilers, 1992; Lynch et al., 1990; Trainor & Hannon, 2012; Trainor & Trehub, 1992).

Despite the studies of metre perception in infancy, little work has examined children's ability to process metrical information. Some timing assessments have been developed for older children and for adults (Gordon, 1980) or have been designed to identify those with serious musical deficits (Peretz et al., 2003). Children's poor motor skills make it difficult to adapt the type of tapping synchronization tasks typically used with adult subjects (Drake, 1993); most involve tapping a single finger, which requires a high degree of fine motor control (Repp, 2005), and measure synchronization over a wide range of tempos over a prolonged period of time (Repp & Su, 2013). When infants and very young children are given the opportunity to move to music, they appear not to synchronize movements with the beat of the music (Eerola et al., 2006; Zentner & Eerola, 2010), indicating that movement synchronization tasks might underestimate young children's ability to perceive synchronization. Drake (1993) investigated factors that influence rhythm reproduction, and found that both Western adults and Western children are better at reproducing rhythms that can be considered simpler in structure. Other work (Einarson & Trainor, subm.; Hannon et al., 2012) has demonstrated that Western children are also better at perceiving timing information in music with simple metric structure, compared to music with complex metric structure.

However, musical timing perception has typically been challenging to measure independent from production or motor synchronization, and few studies have examined beat perception abilities in the absence of an overt motor response in either adults or children (see Repp, 2005, and Repp & Su, 2013, for reviews). Iversen and Patel (2008) developed a Beat Alignment Task (BAT) for adults that measures both perception of beat alignment as well as ability to synchronize to a musical beat and, as such, dissociates overt motor synchrony from perceptual sensitivity. In the perception portion of the BAT, adult subjects judge whether an isochronous auditory tone superimposed over each of 12 musical excerpts is correctly aligned with the beat. Incorrectly aligned beats were either too fast or too slow by 10% of the inter-beat interval, or too early or too late by 25%.

We developed a child-friendly complex Beat Alignment Task (cBAT; Einarson & Trainor, subm.) that can be used with both children and adults. It extends the perception component of the BAT paradigm (Iversen & Patel, 2008) in two notable ways: first, by adding dynamic visual stimuli to accompany short musical excerpts and, second, by including music excerpts with both simple and complex metres. Half of the musical excerpts in the stimulus set are in 4/4 time (simple metric structure), while the other half are in 5/4 or 7/4 time (complex metric structure).

The original cBAT from Einarson and Trainor (subm.) uses dynamic videos of hand puppets that drum along to short musical sequences. An isochronous drumbeat is superimposed on each of the musical tracks such that a woodblock tapping sound matches the drumming movements of the puppets. On each trial, two videos are presented successively, each with the same musical track, but a different puppet. In one video the sound of the puppet's drumming is correctly aligned to the beat of the music; in the other, the puppet drums out of phase (25% too early or too late) or at an incorrect tempo (10% too fast or too slow) relative to the musical excerpt being presented. Participants choose which puppet in each pair is the better drummer. Einarson and Trainor (subm.) found that five-year-old children were sensitive to both phase and tempo misalignments under these conditions, but only in the context of simple metre music. Error detection was not significantly different from chance for complex metre music. As such, we expect that children will show the same perceptual bias for music with culturally typical simple metric structures in the present study.

Using drumming puppets is very engaging and these stimuli kept children attentive through the procedure. However, the stimuli do contain both auditory and visual cues to beat alignment, because trials combine musical examples with video clips of moving puppets. On the one hand, it is possible, though unlikely, that children were relying on visual rather than auditory cues to make their judgments. On the other hand, research has shown that adults perform poorly when attempting to detect misalignment of music and movement in audiovisual tasks (e.g., Phillips-Silver et al., 2011; Vatakis & Spence, 2006a, b). Perception of auditory and visual simultaneity is also vulnerable to the effects of exposure to stimulus asynchrony (Harrar & Harris, 2008), suggesting the perception of synchrony can be modified by experience. Further studies show that visual cues can influence the perception of musical sounds (Schutz & Lipscomb, 2007) and the processing of speech sounds in auditory cortex (Sams et al., 1991).

Most studies of multimodal integration focus on situations where two sensory modalities receive incongruent information about a source (De Gelder & Bertelson, 2003). In addition to those already described, one of the best-known examples of this type of multisensory integration is the McGurk effect (McGurk & MacDonald, 1976), in which there is a mismatch between the presented speech sound and lip movements. In audiovisual tasks designed to promote a perception of simultaneity, there is considerable perceptual tolerance of asynchrony between auditory and visual information (e.g., Sekuler et al., 1997).

It is important to note, however, that in the dynamic cBAT stimuli, the auditory stimulus (isochronous woodblock taps) and visual stimulus (puppet tapping gestures) were synchronous and congruent. As such, there was no misalignment between the auditory and visual modalities. One puppet in every pair simply produced audio-visual drumming that had either incorrect phase or inappropriate tempo relative to the beat of the musical excerpt being presented, leading to mismatched information *within* the auditory stimulus. Even so, it has been established that even congruent visual information can influence perception of simultaneously presented auditory information (Schutz & Lipscomb, 2007) and judgments about musical performance (Tsay, 2013). If auditory perception is, in fact, vulnerable to the effects of simultaneously presented visual information, it is possible that the dynamic visual stimuli could have impaired children's ability to detect beat misalignment in the auditory stimulus.

To address the question of whether young children's sensitivity to the beat as measured in in the cBAT is affected by the presence of dynamic visual information (i.e., moving videos), we tested five-year-old children using either the original cBAT, or a modified version of the cBAT that used static visual images accompanied by the same auditory stimuli.

In addition to administering either the dynamic or static version of the cBAT, we also assessed children's cognitive abilities using standardized measures of memory and vocabulary. Adults' working memory capacity has been found to correlate with their performance on rhythm reproduction tasks (Bailey & Penhune, 2010), and children's ability to discriminate rhythm sequences has been found to correlate with auditory working memory, as measured by a digit span task (Strait et al., 2011). It is also possible that the moving video stimuli would place greater demands on children's working memory than the comparatively impoverished static images, in which case children's memory capacity would contribute to differences in performance between conditions. We examined the relationship between memory span and beat perception by having each child complete the forward and backward components of the digit span subtest from the Wechsler Intelligence Scale for Children (WISC-IV; Wechsler, 2003). Because children needed to understand the instructions in order to perform the cBAT, we also administered the Peabody Picture Vocabulary Test (PPVT-4; Dunn & Dunn, 2007), which is a standardized measure of receptive vocabulary. The PPVT can also serve as a proxy for verbal IQ as it correlates highly with various IQ instruments (Dunn & Dunn, 2007). This allowed us to determine whether beat perception ability might be related to cognitive ability more generally.

The present experiment used two versions of the cBAT to assess whether visual condition (either dynamic or static) would influence five-year-old children's beat alignment sensitivity. We predicted that, as in previous work (Einarson &

Trainor, subm.), children would be sensitive to beat alignment in the context of both tempo- and phase-related errors, but that they would be better able to detect beat alignment in musical excerpts with simple metric structures compared to in musical excerpts with complex metric structures. Additionally, we investigated whether there would be an improvement in children's beat alignment sensitivity when dynamic visual information was removed from the cBAT task, as a result of reduced perceptual inference with the auditory stimuli.

2. Methods

2.1. Participants

The final sample included 84 five-year-old children (41 female, mean age = 5 years, 6 months) who were not enrolled in formal music lessons at the time of testing, and had no previous formal instrumental training. An additional 25 children were excluded from the final analyses because of failure to complete all tasks ($n = 7$), equipment failure ($n = 7$), having formal musical training ($n = 3$), having a diagnosed developmental disorder ($n = 3$), experimenter error ($n = 3$), or vocabulary scores more than two standard deviations below the mean ($n = 2$).

Subjects were assigned to a cBAT condition with either dynamic videos ($n = 48$, 24 female) or static images ($n = 36$, 17 female). Within the dynamic video condition, participants were randomly assigned to either the phase error ($n = 24$, 12 female) or tempo error ($n = 24$, 12 female) condition. In the static image condition, participants were also randomly assigned to either the phase error ($n = 18$, 9 female) or tempo error ($n = 18$, 8 female) condition.

All participants were recruited from the McMaster Infant Studies Group database. Annual family income was measured on a six-point scale (1=<$30,000, 6=>$150,000; Canadian dollars, data missing for nine children). The average family income was between $90,000 and $120,000 per year for children in both conditions. All protocols were approved by the McMaster Research Ethics Board, and informed consent was obtained from all parents.

2.2. Stimuli

2.2.1. Complex Beat Alignment Test

The stimulus set for the *dynamic version* of the cBAT was identical to that used in Einarson and Trainor (subm.). Stimuli consisted of videos of hand puppets drumming along to 12 musical excerpts. Each excerpt ranged in length from 10 to 22 seconds, and tempo ranged from 88 to 172 beats per minute. The excerpts were organized into six pairs that were matched according to genre, tempo, and instrumentation, and all excerpts were selected from existing commercial audio recordings. One excerpt in each pair had simple metric structure (4/4 time signature) common in Western music compositions. The other had a complex metric structure (either a 5/4 or 7/4 time signature). These complex structures are less common in Western music compositions.

In all cases, the visual stimuli were accompanied by a musical excerpt with superimposed isochronous woodblock tapping sounds. The accompanying audio stimuli were created by determining the average tempo of each musical sample using *MixMeister BPM Analyzer* software, and then creating woodblock tapping sounds with the same tempo in *Apple Garageband 2009*. The tapping was aligned with the music by a human listener to create a *correctly aligned* ('on-the-beat') woodblock tapping track for each excerpt, as well as two *phase error* tapping tracks (woodblock taps were phase shifted to be either 25% too early or 25% too late relative to the beat of the music) and two *tempo error* tapping tracks (speed of the woodblock taps was either 10% too fast or 10% too slow relative to the beat of the music).

The standard track, two out-of-phase tracks, and two incorrect-tempo tracks were created for ten of the twelve musical excerpts and were used to create the test trials in both versions of the cBAT. The remaining pair of musical excerpts (one with simple metre and one with complex metre) served as practice trials in the simple metre and complex metre blocks, respectively. Correctly aligned standard tracks were created for these remaining two musical excerpts. Additionally, *randomized* woodblock tracks with pseudo-random non-isochronous taps were created to serve as the “wrong” answer for each of the practice excerpts.

The accompanying visual stimuli were created by filming 24 hand puppets tapping on a toy drum along to the auditory stimuli described above. The two paired puppets on each trial were semantically related (e.g., horse and cow), and each puppet in the pair drummed along to a single simple-metre musical excerpt in the first error condition, and a complex-metre excerpt for the other error condition. Error type within each condition was balanced between subjects so that half of the children in each condition saw a particular puppet make one type of error (e.g., fast or early) and the other half of the children saw that puppet make the other error type (e.g., slow or late).

The result of this balanced design was the creation of six videos for each of the ten musical excerpts that were used in test trials in the dynamic condition (two standard videos, one for each puppet in the pair, two phase shifted error videos from one puppet, and two tempo shifted videos from the other puppet), for a total of 60 test videos, and eight practice videos (four standard and four randomized). In the static condition, there were a total of 68 matched still images accompanied by the same auditory stimuli used in the dynamic condition.

The *static version* of the cBAT was identical to the dynamic version, except that still images of the puppets replaced the videos. Specifically, the images were created by taking a single frame from the beginning of each video from the dynamic condition, and pairing the appropriate audio track with this matched still image. In both conditions, visual stimuli were presented to the child on a 24-inch Samsung SyncMaster 2494 widescreen LCD monitor, and the audio files were played at a comfortable volume, adjusted according to each child’s preference, through M-Audio AV-40 monitor speakers. The experimenter managed the playlists through iTunes 11.0.2(26) on a 21.5-inch Apple iMac (11.2) Core i5.

2.2.2. Digit Span Test

All participants completed a subtest of the Working Memory Index (WMI) from the Wechsler Intelligence Scale for Children, 4th edition (WISC-IV; Wechsler, 2003). This test requires the child to repeat strings of digits produced by the experimenter. In the digits forward section, the child is required to repeat back digits in the same order they are provided. The digits backward section requires the child to repeat the digits in the reverse order from the experimenter. The task begins with a two-digit string, and the number of digits increases by one on every second trial. In each section, testing stops when the child fails to correctly repeat both strings of digits of a given length.

2.2.3. Peabody Picture Vocabulary Test

All participants completed the Peabody Picture Vocabulary Test, 4th edition (PPVT-IV; Dunn & Dunn, 2007), a standardized measure of receptive vocabulary. In this test, the experimenter says a word, and the child chooses which image best represents the vocabulary item from a set of four cartoon pictures. Vocabulary items gradually increase in difficulty. Raw scores were standardized according to the subject’s age, and normalized as described in the Examiner’s Manual.

2.2.4. Parental Questionnaire

While the child was tested, parents completed a questionnaire about their child’s health, language experience, extracurricular activities, and formal or informal music experience. Optional demographic questions regarding parental education and income were answered by only a subset of parents.

2.3. Procedure

Children were tested during one session that was approximately 45 minutes in length. The cBAT (either dynamic or static) was administered in two separate blocks, where one block consisted of the simple metre stimuli and the other the complex metre stimuli. Block order was counterbalanced across children.

Each trial consisted of a pair of two musical excerpts accompanied by drumming. In one, the drumming was correctly aligned to the music. In the other the drumming was either out of phase or out of tempo relative to the excerpt, as described above. In the dynamic cBAT condition, the audio tracks were accompanied by the videos of the puppets drumming whereas in the static cBAT condition, the audio tracks were accompanied by a static image rather than a dynamic video of the puppets drumming. Error type was a between-subjects variable, so that half of the participants were presented with only tempo shift errors for all test trials, and the other half with only phase shift errors. In all four conditions (error-type condition crossed with visual condition), each block consisted of one practice trial (with one puppet whose drumming was correctly aligned, and one that drummed along to the randomized woodblock stimuli), followed by five test trial pairs of puppet stimuli. For each condition, half the children heard two trials on which one puppet was too fast (or early) and three that were too slow (or late); in the other half of children these were reversed. Additionally, half the trials had errors presented by the first puppet in the pair, and half by the second puppet.

Before the cBAT was administered, participants were instructed that they were about to hear two puppets drumming along to the same song, and that they had to determine which puppet would be a better drummer for a band. Before each trial, the experimenter showed the child the two puppets that would be presented in that trial, and identified them by name (for example, “This is my friend the bunny, and this is my friend the dog”). After the child heard both performances, the experimenter placed the two actual puppets in front of the child on the table, and asked the child to place a ribbon on the puppet who should win the prize. The practice trial at the beginning of each block was repeated until the participant correctly identified the puppet whose drumming was correctly aligned to the music. Test trials were not repeated, and the child received no feedback regarding their choices. The order of tasks was the same for all subjects: Block 1 of the cBAT, Digit Span task, Block 2 of the cBAT, followed by the PPVT. At the end of the testing session, each participant received a certificate of participation and a small prize.

3. Results

We first examined children’s ability to detect musical beat misalignment when presented with either dynamic visual movement or static visual images. The dependent measure was the proportion of the time the five-year-old children selected the puppet that played the standard version as the best drummer.

We conducted a 2×2×2 ANOVA with visual condition (dynamic, static) and error type (phase, tempo) as between-subjects factors and metre type (simple, complex) as a within-subjects factor. There was a significant main effect of metre type, $F(1,46) = 14.25$, $p < 0.001$, but no main effect of visual condition, $F(1,46) = 0.23$, $p = 0.63$ and no main effect of error type, $F(1,46) = 0.06$, $p = 0.81$. There were also no significant interactions (all $p > 0.2$). These results suggest that, given identical auditory information, children’s ability to judge beat misalignment is not affected by whether accompanying visual information is static or dynamic (see Fig. 1).

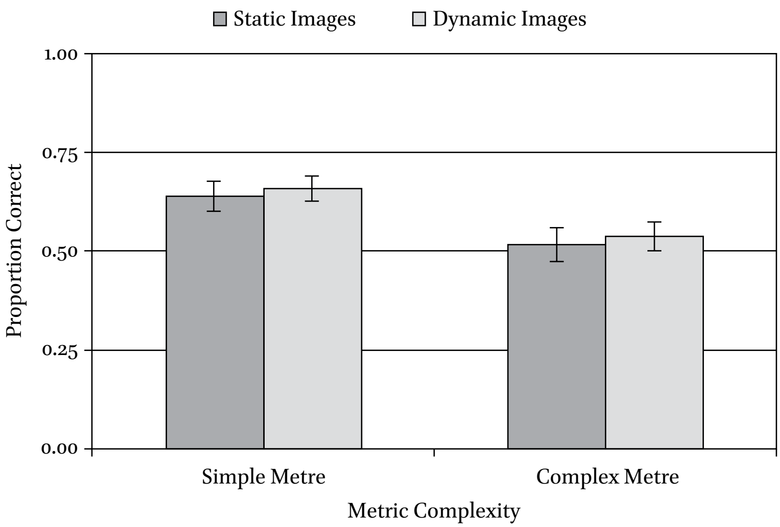

Figure 1. Perception scores from the dynamic and static versions of the cBAT. Because there was no main effect of error type, data have been collapsed across phase and tempo conditions. The proportion of trials on which five-year-old children chose the correctly aligned drumming is not significantly different for static versus dynamic visual conditions (p = 0.63), for either simple or complex metre stimuli, but there is a significant main effect of metre type (p < 0.001). Error bars represent standard error.

Although there was no main effect or interactions involving visual condition, we examined each visual condition separately in order to verify that the effect of metre type was present in both visual conditions, and to determine in which conditions performance was significantly above chance levels. For the dynamic visual condition, a 2×2 ANOVA with error type (phase, tempo) as a between-subjects factor and metre type (simple, complex) as a within-subjects factor revealed only a significant main effect of metre type, $F(1,46) = 6.40$, $p = 0.013$. Children were better at detecting beat misalignment in simple metres (proportion correct = 0.65, SD = 0.22) than complex metres (proportion correct = 0.54, SD =.25) and this was not affected by whether the error was a phase or tempo error. Furthermore, performance was above chance levels for simple meter for both error types (tempo error: $p < 0.001$; phase error: $p = 0.014$) but not for complex meter (tempo error: $p = 0.18$; phase error $p = 0.30$). These results replicate those of Einarson and Trainor (subm.).

For the static visual condition, the 2×2 ANOVA again revealed a significant main effect of metre type, $F(1,34) = 4.66$, $p = 0.038$. As in the dynamic visual condition, children were better at detecting beat misalignment in simple metres (proportion correct = 0.64, SD = 0.23) than complex metres (proportion correct = 0.52, SD = 0.26). Furthermore, performance was above chance levels for simple meter for both error types (tempo error: $p = 0.02$; phase error: $p = 0.005$) but not for complex meter (tempo error: $p = 0.42$; phase error: $p = 0.37$). These results replicate those of Einarson and Trainor (subm.) and suggest that, given identical auditory information, children's ability to judge beat misalignment is not affected by whether an accompanying visual stimulus is static or dynamic.

Overall performance on the cBAT (combined across visual conditions) was not significantly correlated with children's total score on the Digit Span test ($M = 10.63$, $SD = 2.41$), $p = 0.65$. Judgment accuracy for simple metre stimuli was not correlated with forward digit span, $p = 0.50$, or backward digit span, $p = 0.86$. Judgment accuracy for complex metre stimuli was also not correlated with forward digit span, $p = 0.38$, or backward digit span, $p = 0.33$. These results suggest that children's working memory, as indexed by the digit span task, does not relate to performance on the cBAT task.

PPVT standard scores ($M = 117.45$, $SD = 10.44$) did not significantly correlate with the simple metre portion of the cBAT, $p = 0.39$, but were significantly correlated with children's performance on the complex metre portion of the cBAT, $r = 0.28$, $n = 88$, $p = 0.01$.

4. Discussion

We set out to examine the degree to which five-year-old Western children show better perceptual sensitivity for culturally typical simple metre music, compared to culturally atypical complex metres, in the presence or absence of dynamic visual information. We used a perceptual measure of beat sensitivity, the cBAT (Einarson & Trainor, subm.), which assessed children's ability to detect phase and tempo misalignment and included musical excerpts with both simple and complex metres. The original cBAT (Einarson & Trainor, subm.) requires children to watch paired videos of puppets, one of whom drums on time relative to accompanying music and one of whom does not. Children are then asked to judge which puppet drummed better. Drumming errors involved either a tempo error (10% too fast or too slow) or a phase error (25% too early or too late) and in all cases, the movements of the puppets were aligned with the drumming sounds that they made. In the present study, we created a modified version of the cBAT that used static visual images accompanied by the same auditory stimuli.

In both versions of the task, children were able to detect drumming that was out of phase with the music or at the wrong tempo when musical excerpts had simple metric structure. This demonstrates that by the age of five, children have the ability to make explicit judgments about beat alignment, and that given identical auditory information, children's ability to judge beat misalignment is not affected by whether an accompanying visual stimulus is static or dynamic. These findings extend previous literature on children's timing development (Bobin-Bègue & Provasi, 2009; Drake, 1993; Einarson & Trainor, subm.; Kirschner &

Tomasello, 2009; Provasi & Bobin-Bègue, 2003; van Noorden & De Bruyn, 2009; Zentner & Eerola, 2010) by showing that five-year-olds are sensitive to both tempo misalignment and phase misalignment in the context of simple-metered music. Replicating Einarson & Trainor (subm.), for both static and dynamic stimuli, judgment accuracy was significantly worse when musical excerpts had complex metric structures compared to simple metric structures, and performance was not above chance for excerpts with complex metre.

Interestingly, children's performance on the complex metre portion, but not the simple meter portion, of the cBAT in the present study was correlated with receptive vocabulary, as measured by the Peabody Picture Vocabulary Task (Dunn & Dunn, 2007). It is possible that having a larger vocabulary leads to a better ability to comprehend and follow the verbal instructions that accompany the beat perception task, leading to improved task performance. However, this is unlikely, as vocabulary score did not affect performance on the simple meter portion, and the instructions were identical for both portions. Alternately, given that the PPVT correlates highly with other assessments of verbal IQ (Dunn & Dunn, 2007), receptive vocabulary scores may reflect children's cognitive ability more generally. This correlation was not significant in the original cBAT study with five-year-old children (Einarson & Trainor, subm.), so further investigation is needed of the factors that contribute to young children's ability to process complex meters that are not prevalent in their cultural environment.

The development of the cBAT enables further testing of several important questions regarding the development of timing abilities. Because the cBAT is a purely perceptual task, it is now possible to test perception and production separately to determine whether there is a dissociation between these skills in young children. The cBAT also offers a means to test metrical enculturation in young children across cultures. Children from some Balkan regions, for example, grow up being exposed to a much wider variety of musical metres, and it has been observed that they learn these structures easily from a young age (Rice, 1994). However, the only cross-cultural investigation of simple and complex musical metre processing to date tested Bulgarian and Macedonian adults (Hannon & Trehub, 2005). If the better performance of Western children in the present study on musical excerpts with simple compared to complex meter is a function of enculturation, then children who are raised in a culture where complex metre music is widespread should show equal sensitivity to beat alignment in music with simple and complex metres.

In sum, we found that Western five-year-old children can detect misalignments in both the tempo and phase of a beat relative to a musical sample with culturally typical simple metre, and that judgment accuracy is not affected by characteristics of the accompanying visual stimulus; children do equally well when the visual display is static or dynamic. For both types of visual stimuli, children's ability to detect misalignments in both the tempo and phase of a musical beat was at chance levels in the context of less familiar complex metre music. The design of the cBAT enables future investigation of the relation between children's perceptual sensitivity and their ability to synchronize movements to a musical beat. It also provides a means to investigate effects of enculturation on metre sensitivity, for example, by comparing the performance of Western children to children from cultures where complex meters are widespread.

Acknowledgements

This paper was supported by grants to LJT from the Natural Sciences and Engineering Research Council of Canada and the Canadian Institutes of Health Research. We thank the parents and children who participated for their assistance with this research, and we also thank Dave Thompson and Gregory Atkinson for technical assistance.

References

Bailey, J. A., & Penhune, V. B. (2010). Rhythm synchronization performance and auditory working memory in early-and late-trained musicians. *Exp. Brain Res.*, *204*, 91–101.

Bobin-Bègue, A., & Provasi, J. (2009). Tempo discrimination in 3- and 4-year-old children: Performances and threshold. *Curr. Psychol. Lett.*, *2*, 1–13.

De Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends Cogn. Sci.*, *7*, 460–467.

Drake, C. (1993). Reproduction of musical rhythms by children, adult musicians, and adult nonmusicians. *Percept. Psychophys.*, *53*, 25–33.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test fourth edition, Manual.* Minneapolis, MN, USA: Pearson Assessments.

Eerola, T., Luck, G., & Toiviainen, P. (2006). An investigation of pre-schoolers' corporeal synchronization with music. In M. Baroni, A. R. Addessi, R. Caterina, & M. Costa (Eds), *Proceedings of the 9th International Conference on Music Perception & Cognition* (*ICMPC9*), *August 2006, Bologna, Italy* (pp. 472–476). Bologna, Italy: Bononia University Press.

Einarson, K. M., & Trainor, L. J. (subm.). Hearing the beat: Young children's perceptual sensitivity to beat alignment varies according to metric structure. *Music Percept.*

Essens, P. J. (1986). Hierarchical organization of temporal patterns. *Percept. Psychophys.*, *40*, 69–73.

Essens, P. J., & Povel, D. J. (1985). Metrical and nonmetrical representations of temporal patterns. *Percept. Psychophys.*, *37*, 1–7.

Friendly, R. H., Rendall, D., & Trainor, L. J. (2013). Plasticity after perceptual narrowing for voice perception: Reinstating the ability to discriminate monkeys by their voices at 12 months of age. *Front. Auditory Cogn. Neurosci.*, *4*, 1–18.

Friendly, R. H., Rendall, D., & Trainor, L. J. (2014). Learning to differentiate individuals by their voices: Infants' individuation of native- and foreign-species voices. *Dev. Psychobiol.* *56*, 228–237.

Gordon, E. E. (1980). *Manual for the primary measures of music audition and the intermediate measures of music audiation*. Chicago, IL, USA: G.I.A.

Gjerdingen, R. O. (1989). Meter as a mode of attending: A network simulation of attentional rhythmicity in music. *Intégral*, *3*, 67–91.

Hannon, E. E., & Trehub, S. E. (2005). Metrical categories in infancy and adulthood. *Psychol. Sci.*, *16*, 48–55.

Hannon, E. E., & Trainor, L. J. (2007). Music acquisition: Effects of enculturation and formal training on development. *Trends Cogn. Sci.*, *11*, 466–472.

Hannon, E. E., der Nederlanden, C. M. V. B., & Tichko, P. (2012). Effects of perceptual experience on children's and adults' perception of unfamiliar rhythms. *Ann. N. Y. Acad. Sci., 1252*, 92–99.

Harrar, V., & Harris, L. R. (2008). The effect of exposure to asynchronous audio, visual, and tactile stimulus combinations on the perception of simultaneity. *Exp. Brain Res.*, *186*, 517–524.

Iversen, J. R., & Patel, A. D. (2008). The Beat Alignment Test (BAT): Surveying beat processing abilities in the general population. In K. Miyazaki, M. Adachi, Y. Hiraga, Y. Nakajima, & M. Tsuzaki (Eds), *Proceedings of the 10th International Conference on Music Perception & Cognition* (*ICMPC10*)*, August 2008, Sapporo, Japan* (p. 465–468). Adelaide, Australia: Causal Productions.

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychol. Rev.*, *83*, 323.

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy: Evidence of perceptual narrowing. *Psychol. Sci.*, *18*, 1084–1089.

Kirschner, S., & Tomasello, M. (2009). Joint drumming: Social context facilitates synchronization in preschool children. *J. Exp. Child Psychol., 102*, 299–314.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA, USA: MIT Press.

London, J. (2004). *Hearing in time: Psychological aspects of musical meter*. Oxford, UK and New York, NY, USA: Oxford University Press.

Lynch, M. P., & Eilers, R. E. (1992). A study of perceptual development for musical tuning. *Percept. Psychophys., 52,* 599–608.

Lynch, M. P., Eilers, R. E., Oller, D. K., & Urbano, R. C. (1990). Innateness, experience, and music perception. *Psychol. Sci.*, *1*, 272–276.

McGurk, H., & MacDonald, J. (1976) Hearing lips and seeing voices. *Nature 264*(5588): 746–748.

Patel, A. D., Iversen, J. R., Brandon, M., & Saffran, J. (2011). Do infants perceive the beat in music? A new perceptual test. In *Proceedings of the 2011 Biennial Meeting of the Society for Music Perception & Cognition* (*SMPC 2011*)*, August 2011,* Rochester, NY, USA.

Peretz, I., Champod, S., & Hyde, K. (2003) Varieties of musical disorders: The montreal battery of evaluation of amusia. *Ann. N. Y. Acad. Sci.*, *999*, 58–75.

Phillips-Silver, J., Toiviainen, P., Gosselin, N., Piché, O., Nozaradan, S., Palmer, C., & Peretz, I. (2011). Born to dance but beat deaf: A new form of congenital amusia. *Neuropsychologia*, *49*, 961–969.

Provasi, J., & Bobin-Bègue, A. (2003). Spontaneous motor tempo and rhythmical synchronisation in 2½- and 4-year-old children. *Int. J. Behav. Dev.*, *27*, 220–231.

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonom. Bull. Rev.*, *12*, 969–992.

Repp, B. H., & Su, Y. H. (2013). Sensorimotor synchronization: A review of recent research (2006–2012). *Psychonom. Bull. Rev.*, *20*, 403–452.

Rice, T. (1994). *May it fill your soul: Experiencing Bulgarian music*. Chicago, IL, USA: University of Chicago Press.

Sams, M., Aulanko, R., Hämäläinen, M., Hari, R., Lounasmaa, O. V., Lu, S. T., & Simola, J. (1991). Seeing speech: Visual information from lip movements modifies activity in the human auditory cortex. *Neurosci. Lett.*, *127*, 141–145.

Schutz, M., & Lipscomb, S. (2007). Hearing gestures, seeing music – Vision influences perceived tone duration. *Perception, 36*, 888–897.

Sekuler, R., Sekuler, A. B., & Lau, R. (1997). Sound alters visual motion perception. *Nature*, *385*(6614), 308.

Snyder, J. S., Hannon, E. E., Large, E. W., & Christiansen, M. H. (2006). Synchronization and continuation tapping to complex meters. *Music Percept.*, *24*, 135–146.

Soley, G., & Hannon, E. E. (2010). Infants prefer the musical meter of their own culture: A cross-cultural comparison. *Dev. Psychol., 46*, 286–292.

Strait, D. L., Hornickel, J., & Kraus, N. (2011). Subcortical processing of speech regularities underlies reading and music aptitude in children. *Behav. Brain Funct.*, *7*, 1–11.

Trainor, L. J., & Hannon, E. E. (2012). Musical development. In D. Deutsch (Ed.), *The psychology of music* (3rd ed.). San Diego, CA, USA: Academic Press.

Trainor, L. J., & Trehub, S. E. (1992). A comparison of infants' and adults' sensitivity to Western musical structure. *J. Exp. Psychol. Hum. Percept. Perform., 18*, 394–402.

Trehub, S. E. (1976). The discrimination of foreign speech contrasts by infants and adults. *Child Dev., 47*, 466–472.

Tsay, C. J. (2013). Sight over sound in the judgment of music performance. *Proc. Natl Acad. Sci.*, *110*, 14580–14585.

van Noorden, L., & De Bruyn, L. (2009). The development of synchronization skills of children 3 to 11 years old. In P. Toiviainen (Ed.), *Proceedings of ESCOM—7th Triennial Conference of the European Society for the Cognitive Sciences of Music*. Jyväskylä, Finland: University of Jyväskylä.

Vatakis, A., & Spence, C. (2006a). Audiovisual synchrony perception for speech and music assessed using a temporal order judgment task. *Neurosci. Lett.*, *393*, 40–44.

Vatakis, A., & Spence, C. (2006b). Audiovisual synchrony perception for music, speech, and object actions. *Brain Res.*, *1111*, 134–142.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children fourth edition.* San Antonio, TX, USA: The Psychological Corporation.

Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proc. Natl Acad. Sci.*, *106*, 2468–2471.

Zentner, M., & Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proc. Natl Acad. Sci.*, *107*, 5768–5773.