Musical training and auditory global-local precedence

Anne T. Gilman (anne.gilman@gmail.com)
Juniata College Psychology Department, Good Hall, 1700 Moore Street
Huntingdon, PA 16652 USA

Elizabeth Ebbets (lizaebbets@gmail.com)
Juniata College Psychology Department, Good Hall, 1700 Moore Street
Huntingdon, PA 16652 USA

Abstract

People demonstrate a consistent tendency to favor holistic or global processing over processing of local details in many perceptual domains; this tendency is called global-local precedence. Formal musical training is associated with qualitative changes in auditory processing, but the number of years required remains unclear, particularly for any perceptual differences between untrained and minimally trained participants. In this study, participants with zero to over ten years' music training identified the direction of pitch changes in three-sound sequences. Only participants with three or more years' training demonstrated a significant global-local precedence. Individuals without musical training consistently identified the local direction of pitch-change sequence elements better than global pitch changes across each sequence. Although musical training was associated both with greater task accuracy and with global-local precedence, improved accuracy did not explain the musically trained participants' preferential processing of global auditory characteristics.

Keywords: Global-local precedence; musical training; temporal processing; auditory perception.

Perceptual learning and cognition

People's experiences can alter their perceptual processing, even as adults. Perceptual learning thus takes part in core cognitive mechanisms underlying storage of past experiences and their subsequent application to similar but non-identical situations later on. Distinctions between global (or holistic) processing and local (or detail) processing have only recently been examined while taking perceptual learning into account. This connection offers particular interest in the auditory domain, given long-established differences in how trained musicians and musically naïve listeners perceive sounds (Bever & Chiarello, 1974).

Global-local precedence

People who become so engrossed in the details of a complex scene or situation that they fail to notice compelling overall patterns are said to be "missing the forest for the trees". This colloquial expression has a parallel in human perception: decades of vision research has documented a bias or *precedence* towards overall or *global* characteristics of an image rather than its *local* details (Navon, 1977). For example, given an image made up of a dozen K's arranged in the shape of an H, viewers are more likely to describe what they see as an H than as a group of K's. When asked to identify either the big letter or the little component letter in a series of many such composite stimuli, viewers are likely to make fewer errors overall when identifying the big rather than the little letters, thus *global-local precedence* can be identified by greater overall accuracy for global trials. Viewers in such studies also show a pattern in their errors for *incongruent* stimuli such as an H made of K's—one could also create a *congruent* stimulus by arranging K's in the shape of a larger K. Viewers more often mistakenly offer the big letter as an answer on local (little-letter trials) than they err by naming the little letter on global trials, showing an uneven influence—greater for global patterns—of these different processing levels on perception. This uneven influence can also be described as interference by global processing in the local decision. A growing body of evidence supports the overall global-local processing distinction in auditory and other non-visual modalities (Justus & List, 2005; List & Justus, 2007, 2010; List, Justus, Robertson, & Bentin, 2007; Ouimet, Foster, & Hyde, 2012; Sanders & Poeppel, 2007).

Cognitive scientists have found ongoing and robust preferential processing of global over local patterns (Love, Rouder, & Wisniewski, 1999; Ripoll & Marty, 2005) even while some boundary conditions have been delineated (Navon, 2003). This precedence is considered an outcome of normal maturation rather than one of learning. A large body of evidence, summarized by Poirel and colleagues (2011), suggests that young children shift from a local to a global visual focus around the age of 6; Poirel et al. found that lowered grey-matter volume corresponded to this shift, again characterizing global precedence as the healthy adult norm. Visual and auditory findings of differences in holistic versus detailed processing bear on issues ranging from autism (Rondan & Deruelle, 2007) to the acquisition of reading skills (Foxton et al., 2003). For instance, exceptions to global-local precedence for auditory stimuli correspond to autism diagnoses, even though the detail-specific advantages shown by autistic participants were not explained by deficits in global processing as previously thought (Mottron, Peretz, & Ménard, 2000). In all of the above investigations, global-local precedence is taken as an indicator of healthy and fully-developed perception, and only very recently has this tendency been evaluated with respect to prior musical training in a parallel auditory study (Ouimet et al., 2012).

Expertise: Musical training

Auditory perception varies qualitatively—not just in accuracy—depending on learning or expertise. Comparing perception of tones presented to the right or left ear has revealed that hemispheric dominance for this task switches following formal music training (Kellar & Bever, 1980; Bever & Chiarello, 1974). Subsequent studies have shown that musical training changes listeners' perception of auditory patterns, both for musical stimuli themselves (Fujioka, Trainor, Ross, Kakigi, & Pantev, 2004; Foxton, Brown, Chambers, & Griffiths, 2004) and for the pitch changes that characterize fluent speech (Moreno et al., 2009). Musical training also contributes to changes in numerous other cognitive processes, and better understanding of the precise impact of such training on auditory perception may clarify ongoing questions about music and its contribution to formal measures of intelligence and academic achievement (Schellenberg & Peretz, 2007).

The one investigation (other than the present study) of the role of prior musical training in auditory global-local precedence found this precedence more strongly demonstrated among less-trained listeners (Ouimet et al., 2012). Accuracy results were contrasted between expert musicians—with at least seven years' formal training—and novices, where the lower-expertise group could have one or two years of prior training, or none at all. This division is consistent with other key precedents in music perception (Warrier & Zatorre, 2002). However, findings in visual perception, where a mere ten hours' practice with a video game significantly and enduringly changed participants' skills (Green & Bavelier, 2003), suggest that the perceptual performance of participants with minimal amounts of musical training should be contrasted with the performance of those with no training at all.

Evaluating global-local precedence and musical training

The present study reevaluates global auditory bias for possible influences of perceptual expertise, employing the same three-part sweep sequences tested in Sanders and Poeppel's (2007) ERP study to contrast performance of participants with varying levels of musical expertise, including no prior training whatsoever.

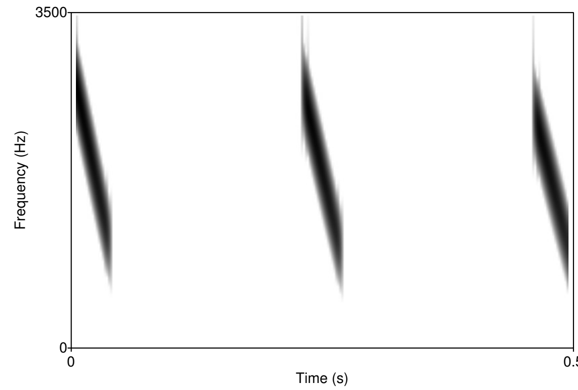
Method

Participants

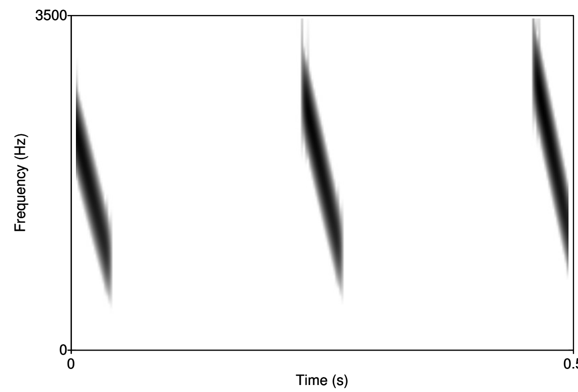
Over multiple semesters, 104 undergraduates with normal or corrected-to-normal hearing participated in the experiment, receiving either course participation credit or snack food as compensation.

Materials

The auditory stimuli selected for this study were created for an ERP study of global versus local auditory processing (Sanders & Poeppel, 2007). Each stimulus consists of three frequency-modulated octave sweeps or chirps that can go up or down in frequency at the global or local level, as follows. Each chirp lasts only 40 ms, with 190 ms of silence separating the chirps; the three-chirp sequence takes 500 ms total. Each chirp travels smoothly up or down one octave—this direction constitutes the *local* direction, and all three chirps of a given sequence go in the same direction. The progression of each sequence is monotonic, with the central pitch of each sweep being higher than the previous one in *global* upward sequences and lower for downward sequences. The difference in central frequency between the first and the last octave sweeps is .2 octaves. In musical terms, if the central frequency of one chirp were a C note, the central frequency of the next chirp would be a bit over a half step higher or lower that that of the first chirp. See Figure 1 for sample spectrograms, obtained using Praat (Boersma, 2001).



(a) A congruent downward stimulus



(b) An incongruent stimulus: local-down/global-up

Figure 1: Stimulus frequency distributions over time.

Congruent stimuli (e.g. Figure 1(a)) have changes in frequency which go in the same direction (up or down) at the local and global levels. Incongruent stimuli (e.g. Figure 1(b)) require the participants to accurately distinguish between levels to answer correctly. The stimulus set includes eleven sequences per condition, namely local-up/global-down, local-up/global-up, local-down/global-down, and local-down/global-up, for a total of 44 stimuli. Frequencies used in these stimuli range between a minimum of 0 Hz and a maximum of 3500 Hz. The intensity of each stimulus file measures between 88 and 92 dB, but participants could control playback volume.

Procedure

Participants were asked to discriminate between upward and downward sweeps (local) and sweep sequences (global) after training involving both passive familiarization and active practice, replicating Sanders & Poeppel's (2007) behavioral protocol. During familiarization, participants first read a description of the type of sounds they were about to hear, e.g. sounds that go down as a whole, and then heard all of those stimuli (global-down ones in this example) in randomized order. During practice trials, the participants heard a sound stimulus while either "PART" (cueing a local judgement) or "WHOLE" (for a global judgement) was displayed at the center of the screen. They indicated by keypress whether the current stimulus was going up or down, and then they received on-screen feedback showing whether or not their answer was correct. After two rounds of practice, participants then enagaged in four blocks of testing trials, which were identical to the practice trials except that no feedback was provided.

After completing the sound discrimination task, participants indicated whether they had ever studied a musical instrument, including voice. They then had the opportunity to clarify if they had studied music for more than one, more than two, or more than five years.

Apparatus

Sounds and instructions were presented on two Macintosh iBook G4s equipped Sennheiser HD 202 headphones and running code written by the first author in PEBL (Mueller, 2006). Participants were free to adjust the computer volume to a comfortable level; they pressed keyboard keys ("U" and "D" for "up" and "down") to record their answers.

Results

An overall global-local precedence in auditory discrimination was demonstrated by our participants, who showed 61.8% accuracy on global trials versus 51.5% accuracy on local trials. A 2x2x2 analysis of variance (ANOVA) comparing the impact of task focus (global or local), stimulus type (congruent or incongruent), and musical training (any or none) showed this difference to be highly significant, $F(1,18296) = 55.8, MSE = 4262, p<.001$. All analyses were performed using R (R Development Core Team, 2005).

As expected, participants were more successful judging the direction of pitch change for congruent stimuli, demonstrating 66.7% accuracy on congruent trials and 51.5% on incongruent trials, $F(1,18296) = 451.8, MSE = 4262, p<.001$.

To evaluate whether our participants demonstrated any global-local precedence, each person's percent accuracy for global and judgements was compared using the simple difference of their average performance in each condition, $M_{global} - M_{local}$. A positive score indicated global-local precedence, while a negative score indicated the opposite. Our participant group as a whole demonstrated an average global-local precedence of .05. This average differed significantly from zero, $t(103) = 3.07, p = .0027$. The significant interaction between task focus and sound type, $F(1,18296) = 82.7, MSE = 4262, p<.001$, further supported our finding of global-local precedence. A Pearson's correlation between this precedence measure calculated over all trials and the same measure calculated over only incongruent trials showed that these incongruent trials were indeed driving the precedence results, $r = .89, t(102) = 20.1, p<.001$. This correlation confirms the presence of the second sign of global-local precedence, where errors on local trials show what can be described as interference from global processing: participants made more errors on local trials, and they did so to a greater extent on incongruent ($M = 0.46$) than congruent ($M = 0.67$) local trials.

Music Training	*Local Advantage*	*Global Advantage*
None	15	5
Some	31	53

Table 1: Number of participants demonstrating each sort of bias according to presence or absence of past musical training

The remaining analyses consistently indicated that musical training is associated with differences in global-local precedence. Those participants who reported any musical training at all showed significantly higher task accuracy (60.2%) than those with none (54.7%), $F(1,18296) = 45.6, MSE = 4262, p<.001$. Reaction times were lower among musically-trained participants. These times were not recorded for half of the participants, due to experimenter error; the available times did not show any interactions with other variables. Greater musical training was associated with significantly different global-local precedence, however, $t(31) = -2.5, p = .017$, with an opposite precedence pattern for those with some musical training ($M = 8\%$) as compared to those with none ($M = -6\%$). This difference clarifies the significant three-way interaction found in the main ANOVA between task focus, sound type, and presence or absence of prior musical training, $F(1,18296) = 21.9, MSE = 4262, p<.001$.

To evaluate this difference without making assumptions about the distributional characteristics of either the training or the precedence score, we compared this indicator of global versus local advantage to reported prior musical training using a χ^2 test (see Table 1). Those with no prior musical background were far more likely to demonstrate better accuracy on local than on global judgements (15 out of 20), while participants with some musical training were more likely (53 out of 84) to demonstrate global-local precedence. This difference was significant, $\chi^2 = 8.022, df = 1, p = .0046$.

To explore this difference further, participants were grouped according to their reported level of musical expertise. Twenty participants reported having no musical training, 16 reported up to one year of training, 17 up to two years, 2 up to five years, and 49 more than five. For these analyses, responses from the two lone participants with between two and five years of training were combined with the most experienced musicians.

As expected, participants with the most musical training in our sample showed the greatest overall accuracy in identifying the direction of pitch change in these three-part auditory sequences: those with more than one year of training showed 61% accuracy; those with up to one year, 57%; and those with none, 55%. All groups performed at above chance levels, $p<.001$.

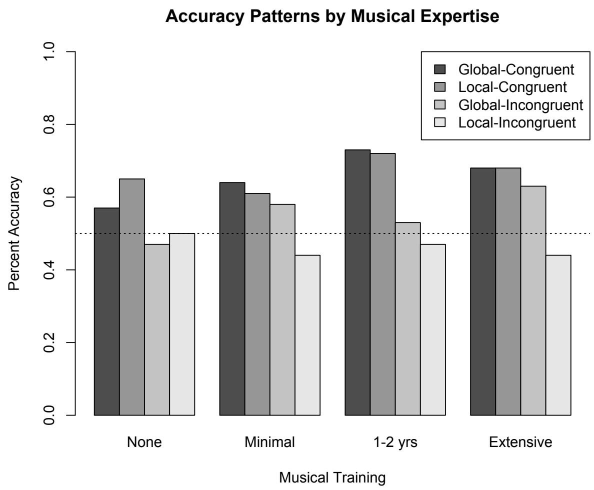


Figure 2: Participant accuracy in identifying direction of pitch changes. The difference between global and local accuracy shows the opposite pattern among non-trained participants compared to the rest both for congruent and incongruent stimuli. Other than incongruent trials for those with 1-2 years' training and local incongruent trials for those with none all accuracies differ significantly from chance perfomance (dotted line).

Discussion

A parallel investigation of global-local precedence in the auditory domain found the opposite results to those presented here, in that their less-trained participants exhibited significant global-local precedence and to a greater extent than their more expert participants (Ouimet et al., 2012). In our study, on the other hand, more expert participants demonstrated global-local precedence, but that precedence only reached statistical significance for those with more than 3 years of training. Those reporting no training at all performed better on local than on global trials, on average. Our study differs from Ouimet et al.'s in multiple ways, however, particularly in the temporal characteristics of the stimuli used and in the grouping of less-expert participants.

Stimulus differences

Ouimet and colleagues used sequences of tones that correspond to notes on a musical scale (Ouimet et al., 2012, p. 2539), based on hierarchical stimuli developed by Justus and List (Justus & List, 2005). Although these tone sequences incorporate a hierarchical structure similar to the sequences of octave sweeps used in our study, there are several important differences.

First, the majority of their sound stimuli lasted much longer. Their three-part sequences of steady tones ranged in duration from 150ms to 600ms, so total stimulus durations ranged from 450ms (comparable to our 500ms, but with no silence) to 1800ms. They found only a "negligible" (Ouimet et al., 2012, p. 2539) relation between stimulus duration and accuracy, however, adding support to List and Justus' arguments for the relational invariance of their tone-sequence tasks (List & Justus, 2010, p.16).

Second, the tone and sweep elements differ sharply in their rate of change of frequency as well as their absolute duration, and the detection of direction of changes in pitch recruits neural resources not required for tone discrimination (Johnsrude, Penhune, & Zatorre, 2000), possibly adding further difficulty for the sweep task. These stimulus differences, however striking, are unlikely to fully explain our differing results, however, since our participants also showed lower accuracy than that found in prior work with these same octave-sweep sequences (Sanders & Poeppel, 2007), while still performing at above-chance levels. Some of our error level may arise from having collected data from pool participants at very different points in the semester (Grimm, Markman, & Maddox, in press). A Pearson's correlation between week of the semester (ranging from 6 to 17, $M = 11.3$) and participant accuracy showed a modest negative correlation, $r = -.25, t(102) = -2.6, p = .0096$. Time of semester showed no interaction with precedence ($p>.5$) or with musical training ($p>.9$); repeating the initial analysis of variance including time of semester did not alter the results obtained.

A third important difference between the two stimulus sets is that in our study, the magnitude of each local change in pitch (one octave) is much greater than each stepped change in the global progression of pitch (one tenth of an octave). In Ouimet et al.'s study (2012), though, global pitch steps are three times the magnitude of local pitch intervals (147 cents, or about 1/8th of an octave). These stimulus differences complicate our interpretation of which accuracy changes are due to global versus local processing and which are due to discrimination difficulty.

Participant differences

The other barrier to explaining the opposite precedence findings between the two studies for less-trained participants lies in the mismatch in the two studies' categorizations of musical training levels, a topic of ongoing debate. While Sanders and Poeppel did not measure musical expertise in their ERP study (2007), other assessments of global-local precedence in the auditory domain did record this participant characteristic. In the study presenting the tone-sequence design discussed above, one experiment had more musically trained than untrained participants, while the second had only musically trained participants (Justus & List, 2005). In one of List and Justus' more recent studies, all participants reported at least 6 years' musical training (List & Justus, 2007), and in another, all participants had at least 2 or 3 years, with each participant group reporting a median of 10 years' training (List & Justus, 2010). Other studies of music perception vary, such as Levitin and colleagues performing fMRI studies on musically-naïve participants who could have up to two years of training (Sridharan, Levitin, & Menon, 2008; Sridharan, Levitin, Chafe, Berger, & Menon, 2007) and other studies where participants all had at least five years' training (Vines, Krumhansl, Wanderley, Dalca, & Levitin, 2011; Vines, Krumhansl, Wanderley, & Levitin, 2006). Ouimet and colleagues (2012) defined musicians as those reporting 7 or more years (versus non-musicians with less than 3 years) of musical training. If most of their non-musicians had 1-2 years' training, then their finding of global-local precedence in that group comes closer to paralleling our findings. If most of their non-musicians had no musical training at all, the contrast between their and our findings would argue all the more strongly for an examination of the stimulus and task differences. None of the precedents listed above assessed differences between those with little training and those with none. This omission may stem in part from widespread reliance on undergraduates as study participants: finding 20 untrained participants required recruiting over 100. However, since a mere ten hours of video game playing can resolve perceptual deficits previously considered to be genetically driven (Feng, Spence, & Pratt, 2007), examining constrasts between untrained and minimally trained listeners is essential.

Refining our examination of perceptual acuity associated with small amounts of musical training increases a risk of possible person confounds in the interpretation of our results. Are individuals with better auditory discrimination more likely to have and/or to continue with music instruction? While this risk cannot be ruled out, our results go beyond what would be predicted from person differences alone. If greater accuracy on our task were purely a matter of better auditory acuity, which may indeed influence individuals' choice to seek or persist in musical training, there would be differences in *overall* accuracy as found here, but *not* an interaction with task focus. Another concern, though, might be that novices' greater difficulty with the overall task might disguise a truly universal global-local precedence that can only be detected among those with greater perceptual acuity. This interpretation was not supported by our data, as shown by a follow-up analysis of each participant group. Under the disguise theory, task accuracy should correlate positively with global-local precedence. As shown in Figure 3, all participant groups showed a zero or, in one case, negative correlation between task accuracy and global-local precedence. In simpler terms, better accuracy with this task did not explain the greater global-local precedence found among more-expert listeners.

There are many ways in which our definition of musical training or expertise could and should be made sharper. Error rates in this data set provide some support for concerns about differing performance across the semester within a participant pool. Even with these concerns, though, our results demonstrate a consistent difference according to musical training in an attribute that is more typically presented as a universal trait in healthy populations. Ongoing scientific interest both in the global-local distinction and in cognitive changes associated with musical training suggests that these results may bear on a wide range of investigations and may specifically help future studies of auditory processing avoid a confound of participant expertise.

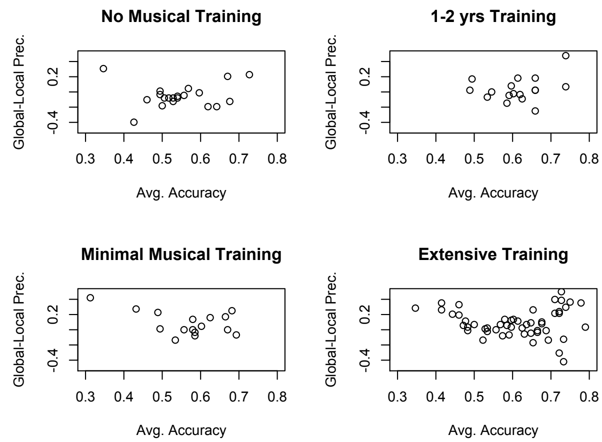


Figure 3: Non-correlation between task accuracy and global-local precedence, assessed separately for each participant group according to their prior musical training.

Acknowledgments

The authors would like to express our gratitude to Lisa Sanders for sharing stimuli and helpful feedback, and to John Limber for improving an earlier draft of this submission.

References

Bever, T. G., & Chiarello, R. J. (1974). Cerebral dominance in musicians and nonmusicians. *Science*, *185*(4150), 537-539.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International*, *5*, 341-345.

Feng, J., Spence, I., & Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychological Science*, *18*(10), 850-855(6).

Foxton, J. M., Brown, A. C., Chambers, S., & Griffiths, T. D. (2004). Training improves acoustic pattern perception. *Current biology*, *14*, 322-325.

Foxton, J. M., Talcott, J., Witton, C., Brace, H., McIntyre, F., & Griffiths, T. D. (2003). Reading skills are related to global, but not local, acoustic pattern perception. *Nature Neuroscience*, *6*, 1097–6256.

Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., & Pantev, C. (2004). Musical training enhances automatic encoding of melodic contour and interval structure. *Journal of cognitive neuroscience*, *16*, 1010–1021.

Green, C. S., & Bavelier, D. (2003). Action video games modify visual selective attention. *Nature*, *423*, 534–537.

Grimm, L. R., Markman, A. B., & Maddox, W. (in press). Semester timing and reward influence performance: End-of-semester syndrome: How situational regulatory fit affects test performance over an academic semester. *Basic and Applied Social Psychology*.

Johnsrude, I., Penhune, V., & Zatorre, R. (2000). Functional specificity in right human auditory cortex for perceiving pitch direction. *Brain*, *123*, 155-163.

Justus, T., & List, A. (2005). Auditory attention to frequency and time: an analogy to visual local-global stimuli. *Cognition*, *98*(1), 31 - 51.

Kellar, L. A., & Bever, T. G. (1980). Hemispheric asymmetries in the perception of musical intervals as a function of musical experience and family handedness background. *Brain and Language*, *10*(1), 24–38.

List, A., & Justus, T. (2007). Auditory priming of frequency and temporal information: Effects of lateralised presentation. *Laterality*, *12*(6), 507 - 535.

List, A., & Justus, T. (2010). Relative priming of temporal local-global levels in auditory hierarchical stimuli. *Attention, Perception, & Psychophysics*, *72*, 193–208.

List, A., Justus, T., Robertson, L. C., & Bentin, S. (2007). A mismatch negativity study of local-global auditory processing. *Brain Research*, *1153*, 122 - 133.

Love, B. C., Rouder, J. N., & Wisniewski, E. J. (1999). A structural account of global and local processing. *Cognitive Psychology*, *38*, 291–316.

Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex*, *19*, 712–723.

Mottron, L., Peretz, I., & Ménard, E. (2000). Local and global processing of music in high-functioning persons with autism: Beyond cerebral coherence? *Journal of Child Psychology and Psychiatry*, *41*, 1057–1065.

Mueller, S. T. (2006). *PEBL: The Psychology Experiment Building Language.* (http://pebl.sourceforge.net/)

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*, 353–383.

Navon, D. (2003). What does a compound letter tell the psychologist's mind? *Acta Psychologica*, *114*, 273–309.

Ouimet, T., Foster, N., & Hyde, K. (2012). Auditory global-local processing: Effects of attention and musical experience. *Journal of the Acoustical Society of America*, *132*, 2536–2544.

Poirel, N., Simon, G., Cassotti, M., Leroux, G., Perchey, G., Lanoë, C., et al. (2011). The shift from local to global visual processing in 6-year-old children is associated with grey matter loss. *PLoS ONE*, *6*(6), e20879.

R Development Core Team. (2005). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Available from `http://www.R-project.org`

Ripoll, T., & Marty, J. (2005). The role of local and global properties in comparison of analogical visual scenes. *Psychonomic Bulletin & Review*, *12*, 178–184.

Rondan, C., & Deruelle, C. (2007). Global and configural visual processing in adults with autism and Asperger syndrome. *Research in developmental disabilities*, *28*, 197–206.

Sanders, L. D., & Poeppel, D. (2007). Local and global auditory processing: Behavioral and ERP evidence. *Neuropsychologia*, *45*(6), 1172 - 1186.

Schellenberg, E. G., & Peretz, I. (2007). Music, language and cognition: unresolved issues. *Trends in Cognitive Sciences*, *12*, 45–46.

Sridharan, D., Levitin, D. J., Chafe, C. H., Berger, J., & Menon, V. (2007). Neural dynamics of event segmentation in music: Converging evidence for dissociable ventral and dorsal networks. *Neuron*, *55*, 521 - 532.

Sridharan, D., Levitin, D. J., & Menon, V. (2008). A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 12569 - 12574.

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., Dalca, I. M., & Levitin, D. J. (2011). Music to my eyes: Cross-modal interactions in the perception of emotions in musical performance. *Cognition*, *118*, 157 - 170.

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., & Levitin, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition*, *101*, 80 - 113.

Warrier, C., & Zatorre, R. (2002). Influence of tonal context and timbral variation on perception of pitch. *Perception and Psychophysics*, *64*, 198–207.