2019

Consonant and dissonant music chords improve visual attention capture

Graham Spurrier

Recommended Citation

Spurrier, Graham, "Consonant and dissonant music chords improve visual attention capture" (2019). *CMC Senior Theses*. 2125.
https://scholarship.claremont.edu/cmc_theses/2125

Claremont McKenna College

Consonant and dissonant music chords improve visual attention capture

Submitted to
Professor Alison Harris

by
Graham H. Spurrier

For
Senior Thesis
Spring 2019
April 29, 2019

Consonant and dissonant music chords improve visual attention capture

Graham H. Spurrier
Claremont McKenna College

Acknowledgments

I wish to express my sincerest gratitude to the Claremont McKenna College psychology department. Every professor I have had the privilege of working with has contributed to making my four years at CMC incredibly memorable. Further, each has demonstrated incredible insight into their respective field, which has encouraged exploration of and interest in many different areas of psychology.

However, I feel it is necessary to highlight a few professors who have been especially impactful in my journey at Claremont McKenna College, whose mentorships have been instrumental in my successes as a student. In addition to being an exceptionally engaging teacher and helpful academic advisor, Professor Wei-Chin Hwang has provided immeasurable guidance and support during these four years.

I have had the privilege of taking three classes taught by Professor Gabriel Cook as well as the opportunity of assisting in his research as an undergraduate. He is responsible for stimulating my interest in cognitive psychology and eventually, music cognition, which prompted the undertaking of this project. Professor Cook has also been incredible to me as a person, especially in particularly rough times.

I have only had the pleasure of knowing Professor Sharda Umanath since last spring, but she has been instrumental in furthering my knowledge of research methods and experimental design. Almost exactly one year ago, we were both anxious about the "viability" of my music tempo study's data. Throughout these last few semesters at CMC, Professor Umanath has been incredibly supportive of my endeavors as a mentor and friend.

Lastly, I would like to express my deepest gratitude to Professor Alison Harris. As my thesis reader, Professor Harris has had to deal with an "adventurous" year: from serial topic uncertainty to diving into a complicated and conflicting area of research, she has handled the ups and downs of my project with grace, while being immensely helpful and instructive in the process. Thank you for all the guidance you have offered.

Acknowledgments would be incomplete without thanking my parents and brother. Thank you for supporting me through the last twenty-three years of my life, and in particular, the more difficult times in these last few years.

Table of Contents

Section

Abstract

Recent research has suggested that music may enhance or reduce cognitive interference, depending on whether it is tonally consonant or dissonant. Tonal consonance is often described as being pleasant and agreeable, while tonal dissonance is often described as being unpleasant and harsh. However, the exact cognitive mechanisms underlying these effects remain unclear. We hypothesize that tonal dissonance may increase cognitive interference through its effects on attentional cueing. We predict that (a) consonant musical chords are attentionally demanding, but (b) dissonant musical chords are more attentionally demanding than consonant musical chords. Using a Posner cueing task, a standard measure of attention capture, we measured the differential effects of consonant chords, dissonant chords, and no music on attentional cueing. Musical chords were presented binaurally at the same time as a visual cue which correctly predicted the spatial location of a subsequent target in 80% of trials. As in previous studies, valid cues led to faster response times (RTs) compared to invalid cues; however, contrary to our predictions, both consonant and dissonant music chords produced *faster* RTs compared to the no music condition. Although inconsistent with our hypotheses, these results support previous research on cross-modal cueing, which suggests that non-predictive auditory cues enhance the effectiveness of visual cues. Our study further demonstrates that this effect is not influenced by auditory qualities such as tonal consonance and dissonance, suggesting that previously reported cognitive interference effects for tonal dissonance may depend on high-level changes in mood and arousal.

Keywords: music, attention, tonal consonance and dissonance, Posner cueing task

Introduction

Music has been a prominent feature in human cultures throughout history, with evidence of musical instruments dating back to 40,000 years ago (Conard, Malina, & Münzel, 2009). Yet the question of why music is prevalent in human cultures has been debated since Aristotle, who believed that music was one of the world's most important unsolved mysteries (Aristotle, trans. 1995). Philosophers, scientists, and psychologists have proposed many different theories of music's function in human society. Darwin believed that music's function was to express emotions, especially in the context of sexual selection (Darwin, 1871). Others have argued that music may serve to promote cooperative, prosocial behavior in group settings (Wallin, Merker, & Brown, 2001). Research on the role of music in human cognition has reported improvements in working memory, intelligence and visuospatial processing, though many of these effects appear to be driven by changes in mood and arousal (Schellenberg, 2005). On the other hand, acquired amusia, or loss of music perception following brain damage, is associated with general deficits in a variety of cognitive domains, suggesting that music perception is linked to other high-level functions (Särkämö, Tervaniemi, Soinila, Autti, Silvennoinen, Laine, & Hietanen, 2009).

Recent research has further distinguished the effects of music on cognition with respect to tonal consonance and dissonance (Masataka and Perlovsky, 2013; Komeilipoor, Rodger, Craig, & Cesari, 2014), a commonly described feature of musical intervals in the Western musical tradition. An interval is a combination of two musical notes, which can be played simultaneously or separately. Whereas Western listeners generally associate consonance with pleasantness and stability, dissonance is associated with harshness and instability, leading to preferences for consonant music (Guthrie & Morrill, 1928; Plomp & Levelt, 1965; McDermott, 2010), though these preferences may depend on enculturation (McDermott, Schultz, Undurraga, & Godoy, 2016). Musical intervals are often viewed along a gradient ranging from most consonant to most dissonant (Schoenberg, 1975), with certain intervals consistently ranked as consonant (e.g. octave, perfect fifth) and other intervals as dissonant (e.g. minor second; Malmberg, 1918; Guthrie & Morrill, 1928). The exact physical qualities that produce consonance or dissonance have been argued for thousands of years. One feature that has been specifically implicated in dissonance is the ratio of note frequencies (Hz) in music intervals and the resulting "beating" of desynchronized harmonics (Helmholtz, 1877; though see Plomp & Levelt, 1965). Another related acoustic quality is roughness, a perceptual quality of rapidly modulated sound, which is a key component of both natural and artificial alarm signals (Arnal, Flinker, Kleinschmidt, Giraud, & Poeppel, 2015).

Research has demonstrated that dissonant sounds are processed more rapidly in consonant sound contexts than vice versa (Schellenberg & Trehub, 1994). This effect is associated with differential amplitude of the mismatch negativity (MMN; Bojorque, Monte-Ordoño, & Toro, 2018), a pre-attentive auditory event-related potential (ERP) that measures perceptual detection of acoustic changes in the immediate environment (Näätänen, 2003). Therefore, Bojorque and colleagues (2018) suggest that consonant and dissonant sounds are perceived differently even at the earliest stages of auditory processing. In a movement synchronization task, participants showed better synchronization performance after listening to consonant but not dissonant metronomes, suggesting that consonant input may enhance motor coordination (Komeilipoor et al., 2014).

Consonance and dissonance have also been argued to affect task performance at the cognitive level. Notably, Masataka and Perlovsky (2013) reported that, relative to participants not hearing music, participants listening to a predominantly consonant Mozart minuet showed facilitation on a Stroop task, with fewer errors and faster response times (RTs) when naming incongruent font colors of printed color words (e.g., “red” printed in blue ink). In contrast, participants exposed to a variant of the same minuet consisting of predominantly dissonant intervals showed increased interference, becoming slower and making more errors for incongruent font colors. Based on these results, Masataka and Perlovsky (2013) proposed that consonant music mitigates cognitive dissonance, thereby reducing cognitive interference in complex cognitive tasks, while dissonant music strengthens cognitive interference. These results join previous data indicating that the presence of background noise and music may distract individuals and worsen performance on the Stroop task, particularly for music with high arousal potential and negative affect (Cassidy & MacDonald, 2007). Together, these studies support the idea that dissonance, in particular, is associated with interference on cognitive tasks.

However, the cognitive mechanism underlying this strengthened interference effect remains unclear. Although Masataka and Perlovsky (2013) interpret the Stroop effect as measuring cognitive dissonance, the specific cognitive functions and systems targeted by the Stroop task continue to be a topic of debate, with explanations including automaticity of reading, speed of processing, and parallel processing of relevant and irrelevant dimensions (MacLeod, 1991). Furthermore, this study used a continuous musical piece, so the onset of consonant and dissonant intervals was not precisely controlled over the course of the experiment (Masataka & Perlovsky, 2013), which complicates precise inference regarding the cognitive origins of the observed interference effect.

One potential auditory factor could be roughness, which is associated with dissonant intervals (Helmholtz, 1877). A recent study demonstrated that roughness facilitates greater speed and accuracy of spatial localization (Arnal et al., 2015). Along with the aforementioned advantage in detecting dissonant stimuli within a consonant context (Schellenberg & Trehub, 1994; Bojorque et al., 2018), these findings suggest that dissonant sounds are more attentionally salient than consonant sounds. Stimuli with high feature salience are believed to capture attention more effectively than low feature salience stimuli, leading to greater interference from high-salience distractors in visual tasks (Mounts, 2005; Zehetleitner, Koch, & Müller, 2013). Thus, although no studies have directly examined the attentional salience of tonal dissonance, these data suggest that dissonant music may produce cognitive interference on the Stroop task through increases in early attentional capture.

To determine if consonance and dissonance have differential effects on attention capture, this study will measure the effects of simultaneously-presented consonant and dissonant music chords during the Posner cueing paradigm (Posner, 1980), a widely-used measure of attention shifting (for a review, see Hayward & Ristic, 2013). In the traditional Posner cueing paradigm (Figure 1), the participant is required to maintain central fixation while monitoring two boxes to the left and right for the appearance of a *target* stimulus (e.g., red circle). When the target appears, the participant is instructed to respond as quickly and accurately as possible with a left or right keypress to indicate the target's location. The target is preceded by a *cue* briefly highlighting one of the possible two locations. The cue may be valid, correctly indicating the location of the target stimulus, or invalid, in the opposite location. Previous research has demonstrated that when the cue is generally predictive of target location, RTs are faster and responses are more accurate following valid cues, whereas invalid cues produce slower RTs and greater error rates. These effects are believed to index attentional capture due to increased weighting of the location associated with the predictive cue (Eckstein, Shimozaki, & Abbey, 2002).

In this experiment, the appearance of the cue was synchronized with a consonant chord, dissonant chord, or no music. Given previous findings of cognitive interference from dissonant intervals (Cassidy & MacDonald, 2007; Masataka & Perlovsky, 2013), we hypothesized that RTs would be slowest for dissonant trials since they would divert attention from the task (Mounts, 2005; Zehetleitner, Koch, & Müller, 2013). Because consonant chords would also constitute an additional environmental stimulus (Cassidy & MacDonald, 2007), we predicted that they would divert attention from the task but to a lesser degree. Thus, RT was expected to be slower for consonant chords than for no music trials, but better than dissonant chord trials. Accuracy was also measured but due to the simplicity of the Posner cueing task, we expected a ceiling effect.

Methods

Participants

The participants in this study were 24 undergraduate students (ages 18-25, 58% female) at a Southern California college. All participants were currently enrolled in a lower-division psychology course at the time of participation. For their participation in the study, participants were awarded research credit. Due to previous research suggesting that differences in perception of consonance and dissonance depend on musical training, participants were also asked whether they had 10 years or more (Rodrigues, Loureiro, & Caramelli, 2013; Bojorque et al., 2018) of musical training (yes/no).

Materials

Although most studies of consonance and dissonance use simple intervals (i.e. one interval presentation at a time; Malmberg, 1918; Guthrie & Morrill, 1928; Schellenberg & Trehub, 1994; Bojorque et al., 2018), here we employed music chords. Chords consist of at least three intervals: the first and second notes, the second and third notes, and the first and third notes. Major chords, which consist of a major third, minor third, and perfect fifth intervals, are consistently rated as consonant by both musicians and non-musicians (Roberts, 1986). A tone cluster is a chord that is constructed by simultaneously playing two or more adjacent semitones (i.e. minor second intervals) at the same time. Tone clusters are associated with extreme dissonance given the presence of multiple minor second intervals (Seachrist, 2003). As the number of dissonant intervals in a chord increases, the more dissonant the chord should become, since the interval ratios become more complicated and the more harmonics there are to beat with one another. Thus, this study uses a C-major chord (i.e. C4, E4, G4) and a tone cluster (i.e. C4, D♭4, D4) as the consonant and dissonant music chords, respectively. These chords are also compatible with the role of familiarity on consonance and dissonance evaluations; the relative prevalence of major chords and tone clusters in popular Western music strongly favors the former (Vos and Troost, 1989).

Design

The experiment used a 2x3 within-subjects design. The independent variables were cue type (valid, invalid) and tone condition (consonant, dissonant, no music). The dependent variable was response time. For a secondary analysis, accuracy was also collected as a dependent variable.

Procedure

This study used a modified version of the Posner cueing task, which is a widely-used measure of spatial attention (Posner, 1980; Hayward & Ristic, 2013). In the standard Posner cueing paradigm, a visual cue is used to attract participants' attention to a location that may present a response target. Here, a similar paradigm was used, but a non-predictive audio sound (i.e. consonant or dissonant chord) could be played with the visual cue (Figure 1). Specifically, from the cue phase until a response as made in the stimulus-response phase, (a) a consonant chord played (i.e. C-major chord), (b) a dissonant chord played (i.e. tone cluster), or (c) no musical chord played. In 80% of the trials, the visual cue correctly predicted target location (i.e. *valid* trial), and in 20% of the trials, the visual cue incorrectly predicted target location (i.e. *invalid* trial). Participants were instructed to respond by pressing a key with their right hand or left hand depending on the location of the target (Appendix A).

Musical chords were presented binaurally via headphones (Panasonic Headphones RP-HT161-K) so as to not provide any auditory cues regarding the spatial localization of the target. Participants had a maximum of 500 ms to respond. Following the cueing and stimulus-response phase, there was an inter-trial interval of 1200-1600 ms. Before beginning the task, participants received a short block of practice trials including feedback about overly slow RTs; however, no feedback was provided during the task itself. There were a total of 450 experimental trials, so to allow participants time to rest, the trials were divided into 15 test blocks, with short breaks in between each block. In total, the experiment took approximately twenty minutes.

After participants had completed the experimental program, they were given a short questionnaire (Appendix B) to fill out. The consent form and questionnaire were administered via iPad.

Results

In this experiment, we examined whether consonance and dissonance produce differences in attentional capture as an explanation for the observation that tonal dissonance increases cognitive interference (Masataka & Perlovsky, 2013). Using the standard Posner cueing paradigm, we measured differences in RT and accuracy of target detection following valid or invalid cues. Cues were presented with one of three tone conditions: No Tone, Consonant, Dissonant. Additionally, musical training (≤10 years, >10 years) was included as a covariate of no interest, as extensive musical training has previously been shown to influence visual attention ability (Rodrigues et al., 2013) and affect perception of consonant and dissonant chords (Bojorque et al., 2018). Thus, RT and accuracy data for each participant were entered into two separate 2 (cue type: valid, invalid) x 3 (tone condition: consonant, dissonant, none) repeated-measure analyses of covariance (ANCOVA). Where necessary, all reported statistics were corrected for violations of sphericity using the Greenhouse-Geisser epsilon.

First, we examined whether we successfully replicated the Posner cueing advantage for valid cue presentations (Figure 2A). Consistent with the previous literature, we found that valid cues were associated with faster RTs (Figure 2A) and higher accuracy (Figure 2B). These results were confirmed by significant main effects of validity for RT ($F(1,22) = 91.02$, $p = 2.82 \times 10^{-9}$, $\eta_p^2 = 0.805$) and accuracy ($F(1,22) = 14.93$, $p = 8.40 \times 10^{-5}$, $\eta_p^2 = .805$). However, the interaction of cue validity with musical training failed to reach significance either for RT ($F(1,22) = 0.125$, $p = 0.73$) and accuracy ($F(1,22) = 2.49$, $p = 0.13$).

Second, we looked at the effect of simultaneous presentation of consonant or dissonant musical chords on target detection performance. Our hypothesis was that musical chords are attentionally demanding and would distract from the Posner cueing task. Thus, we predicted that RT would be slower when musical chords were present when compared to a No Music control condition, and this effect would be greatest in the Dissonant chord condition. Alternatively, another proposed outcome was that the Consonant chord condition would reduce cue-target interference (leading to faster RTs), and Dissonant chords would increase cue-target interference (producing slower RTs) in comparison to the No Tone control condition (Masataka & Perlovsky, 2013).

Although there was no significant difference in accuracy between tone conditions ($F(1.55, 34.1) = 1.11$, $p = 0.32$), we found a significant main effect of tone condition on

RT ($F(2, 44) = 20.2$, $p = 6.12 \times 10^{-7}$, $\eta_p^2 = 0.478$). As shown in Figure 2, this effect was driven by ***faster*** RTs for ***both*** Consonant and Dissonant chord conditions relative to the No Tone control condition (No Tone: $M = 0.303 \pm 0.03$ s; Consonant: $M = 0.289 \pm 0.031$ s; Dissonant: $M = 0.288 \pm 0.032$ s). The interaction of Cue Validity and Tone Condition failed to reach significance for both accuracy ($F(1.52,33.4) = 1.38$, $p = 0.26$) and RT ($F(1.52,33.5) = 0.38$, $p = 0.63$, $\eta_p^2 = 0.02$). Additionally, all interactions of Tone Condition with the covariate of musical training were also non-significant (Fs < 1). These results suggest that the simultaneous presentation of an audio cue with a visual target facilitates faster reaction time to that target, but the tonal quality (consonant or dissonant) of the audio cue is irrelevant in these early stages of perceptual processing.

Because RT data can become skewed when introducing decision-making or interference procedures into a task (Luce, 1986), we also inverse-transformed the RT by taking its reciprocal, a manipulation that is known to normalize the RT distribution and reduce the effects of slow outliers (Whelan, Effective Analysis of Reaction Time Data, 2008). However, we found that inverse-RT effects were highly similar to those for raw RT, and thus are not reported here.

Discussion

This study's hypotheses were that (a) the introduction of musical chords would decrease performance compared to no music and (b) dissonant chords would decrease performance more than consonant chords on the Posner cueing paradigm. Results showed that RTs were faster during experimental trials when consonant and dissonant chords were played, with no differential effects of consonance vs. dissonance on performance.

Thus, the experiment supported neither of the study's hypotheses but nonetheless provided interesting results.

The finding that the introduction of novel auditory cues into the Posner cueing paradigm improved performance is consistent with previous research on cross-modal attention (Spence & Driver, 1996; Spence & Driver, 2004; Hancock, Oron-Gilad, & Szalma, 2007). Studies of cross-modal attention examine the differential effects of presenting stimuli in different modalities (e.g., visual, auditory, tactile). In the Masataka & Perlovsky study (2013), experimental stimuli are cross-modal: the Stroop is presented visually, while the music conditions are auditory, meaning the stimuli appeal to different sensory modalities. Similarly, this study also paired visual stimulus cues with auditory stimuli (i.e. consonant and dissonant chords), but unlike in Masataka & Perlovsky (2013) the onset of consonance and dissonance was precisely controlled throughout the experiment.

This study's results suggest that the simultaneous presentation of the auditory stimuli and visual cues facilitated "multi-modal integration", which occurs when the presentation of auditory and visual stimuli are spatially congruent and close in temporal proximity (see Calvert, 1998). This perceptual integration process has been shown to increase detectability of visual targets in experiments using cross-modal stimuli (Vroomen & Gelder, 2000; Teder-Sälejärvi, Di Russo, McDonald, & Hillyard, 2005; Lippert, Logothetis & Kayser, 2007) and improve visual performance (e.g. RT, accuracy, spatial judgments) in experiments using cross-modal cuing (Spence & Driver, 1997; Spence & Driver, 1998; McDonald, Teder-Sälejärvi, & Hillyard, 2000). The attentional benefits associated with cross-modal cueing begin to dissipate with higher stimulus onset asynchronies, a phenomenon known as inhibition of return (IOR; see Klein, 2000 for a review), but in this study, audio and visual stimuli were presented simultaneously, which suggests the compatibility of this study's design with cross-modal cueing effects. Our results support previous research, as the sheer presence of a simultaneous, non-predictive auditory stimulus improved RT.

At the same time, our data suggest that early multi-modal integration does not depend on the acoustic qualities of the auditory stimulus, as both consonant and dissonant chords were equally effective in improving attentional capture. These findings also raise questions about the privileged status assigned to acoustic qualities such as dissonance and roughness (Arnal et al., 2015). At least at the early stages of attentional capture tested here, we found no difference in spatial localization of a simultaneous visual cue. However, future research should examine attentional capture by consonance and dissonance when the only cue for the visual target is auditory, rather than a combination of visual and auditory cues. Another possible direction would be to determine the effect of having consonant and dissonant sounds precede the onset of the visual cue, which might facilitate performance deficits associated with IOR. In terms of the predictability of audio cues, further research could introduce directionality into the audio stimulus to determine if this performance facilitation effect occurs only with non-predictive cues.

Our failure to find differential effects of consonant and dissonant chords is also inconsistent with previous research on the cognitive effects of dissonance (Cassidy & MacDonald, 2007; Masataka & Perlovsky, 2013). One possibility is that the Posner cueing task is too simple to observe an effect on attention. Previous research examining differential effects of music (e.g. music intensity, affect, consonance and dissonance, genre) on attentional performance has relied on more complex tasks as measures of attention, like reading comprehension, memorization, and simulated driving tasks (Beh & Hirst, 1999; Cassidy & MacDonald, 2007; Thompson, Schellenberg, & Letnic, 2011; Masataka & Perlovsky, 2013). Thus, it is possible that the differences in consonance and dissonance observed in Masataka and Perlovsky (2013) are only present in more complex, attentionally-demanding tasks, consistent with the idea that short-term effects of music on cognitive function arise from changes in arousal and mood (Schellenberg, 2005).

Another factor worth exploring is the relative familiarity of consonance and dissonance. Researchers have noted that consonance is more prevalent than dissonance in Western music (Vos & Troost, 1989) and that preference for consonance music is dependent on exposure to Western music (McDermott, Schultz, Undurraga, & Godoy, 2016). Thus, the amount of exposure to dissonant music could moderate the relationship between dissonance and attention capture such that increased exposure to dissonant sounds might mitigate the attentional effects of tonal dissonance. In this study, consonant and dissonant chords appeared with equal frequency, which could have familiarized the participants with the dissonant chords. In other studies, processing advantages for dissonant sounds have been reported in consonant sound contexts (Schellenberg & Trehub, 1994; Bojorque et al., 2018), which suggests that dissonant sounds may need to appear infrequently (i.e. violate tonal expectations) in order to capture more attention than consonant sounds. To test this idea, a replication could be conducted with an oddball paradigm (as first discussed in Squires, Squires, & Hillyard, 1975), comparing the effects of (a) a repeated consonant chord with a different consonant chord as the oddball and (b) a repeated consonant chord with a different dissonant chord as the oddball. Given the proposed role of familiarity in violating tonal expectations and facilitating increased attention capture by tonal dissonance, we would expect to see more rapid processing for (a) than (b).

Moreover, attentional demand differences between tonal consonance and dissonance may be sensitive to further expectations of chord relationships. In this experiment, the consonant and dissonant chords were presented independently of one another with an inter-trial interval (ITI) of 1200-1600 milliseconds in between each presentation. Masataka and Perlovsky (2013) manipulated continuous harmonic relationships by modifying certain consonant intervals in a minuet to be more dissonant. This should produce independent chord dissonance (i.e. vertical dissonance) and also introduce dissonance into harmonic relationships (i.e. horizontal dissonance) that would otherwise be consonant. By not resolving consonant chord progressions, the dissonant musical piece is again violating expectations that have been developed in Western music. The formation and violation of harmonic expectations (e.g. chord progressions) have been noted to play a role in perception (see Schmuckler, 1989). Thus, the dissonance manipulation in our study could have been weaker than previous studies that also took continuous, harmonic relationships into account. As a priority, future research should introduce both vertical (i.e. interval, chord structure) and horizontal dissonance (e.g. harmonic relationships, chord progressions) since effects associated with tonal dissonance will be theoretically strengthened: a replication of this study could be conducted where the single consonant or dissonant chords are replaced by consonant or dissonant chord progressions in order to strengthen tonal dissonance.

In conclusion, this study explored whether the previously described cognitive effects of tonal consonance and dissonance are explained by differential attentional capture. Although we failed to find any evidence for additional distraction by dissonant chords, our results support previous findings that cues with appeal to multiple sensory modalities have processing advantages. Whereas tonal consonance and dissonance do not appear to be differentiated in the early stages of attentional capture, cognitive effects related to these acoustic qualities may instead arise from high-level changes in mood and arousal.

References

Arnal, L. H., Flinker, A., Kleinschmidt, A., Giraud, A.-L., & Poeppel, D. (2015). Human screams occupy a privileged niche in the communication soundscape. *Current Biology*, *25*(15), 2051–2056.

Barnes, J., & others. (1995). *The Cambridge Companion to Aristotle*. Cambridge University Press.

Beh, H. C., & Hirst, R. (1999). Performance on driving-related tasks during music. *Ergonomics*, *42*(8), 1087–1098.

Calvert, G. A. (2001). Crossmodal processing in the human brain: insights from functional neuroimaging studies. *Cerebral Cortex (New York, N.Y. : 1991)*, *11*(12), 1110–1123.

Cassidy, G., & MacDonald, R. A. (2007). The effect of background music and background noise on the task performance of introverts and extraverts. *Psychology of Music*, *35*(3), 517–537.

Conard, N. J., Malina, M., & Münzel, S. C. (2009). New flutes document the earliest musical tradition in southwestern Germany. *Nature*, *460*, 737.

Crespo-Bojorque, P., Monte-Ordoño, J., & Toro, J. M. (2018). Early neural responses underlie advantages for consonance over dissonance. *Neuropsychologia*, *117*, 188–198.

Darwin, C. (1871). *The descent of man, and selection in relation to sex.* (1st ed., Vol. 1). London: John Murray.

Driver, J., & Spence, C. (1998). Cross–modal links in spatial attention. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *353*(1373), 1319–1331.

Eckstein, M. P., Shimozaki, S. S., & Abbey, C. K. (2002). The footprints of visual attention in the Posner cueing paradigm revealed by classification images. *Journal of Vision*, *2*(1), 3–3.

Guthrie, E., & Morrill, H. (1928). The fusion of non-musical intervals. *The American Journal of Psychology*.

Hancock, P. A., Oron-Gilad, T., & Szalma, J. L. (2007). Elaborations of the multiple-resource theory of attention. *Attention: From Theory to Practice*, 45–56.

Hayward, D. A., & Ristic, J. (2013). Measuring attention using the Posner cuing paradigm: the role of across and within trial target probabilities. *Frontiers in Human Neuroscience*, *7*, 205.

Helmholtz, H. (1877). On the sensations of tone. *Trans. AJ Ellis, Dover, New York*.

Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences*, *4*(4), 138–147.

Komeilipoor, N., Rodger, M. W. M., Craig, C. M., & Cesari, P. (2015). (Dis-)Harmony in movement: effects of musical dissonance on movement timing and form. *Experimental Brain Research*, *233*(5), 1585–1595.

Lippert, M., Logothetis, N. K., & Kayser, C. (2007). Improvement of visual contrast detection by a simultaneous sound. *Brain Research*, *1173*, 102–109.

Luce, R. D., & others. (1986). *Response times: Their role in inferring elementary mental organization*. Oxford University Press on Demand.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychological Bulletin*, *109*(2), 163.

Malmberg, C. F. (1918). The perception of consonance and dissonance. *Psychological Monographs*, *25*(2), 93.

Masataka, N., & Perlovsky, L. (2013). Cognitive interference can be mitigated by consonant music and facilitated by dissonant music. *Scientific Reports*, *3*, 2028.

McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2010). Individual differences reveal the basis of consonance. *Current Biology : CB*, *20*(11), 1035–1041.

McDermott, J. H., Schultz, A. F., Undurraga, E. A., & Godoy, R. A. (2016). Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature*, *535*(7613), 547.

McDonald, J. J., Teder-Salejarvi, W. A., & Hillyard, S. A. (2000). Involuntary orienting to sound improves visual perception. *Nature*, *407*(6806), 906–908.

Mounts, J. R. (2005). Attentional selection: A salience-based competition for representation. *Perception & Psychophysics*, *67*(7), 1190–1198.

Näätänen, R. (2003). Mismatch negativity: clinical research and possible applications. *International Journal of Psychophysiology*, *48*(2), 179–188.

Plomp, R., & Levelt, W. J. M. (1965). Tonal consonance and critical bandwidth. *The Journal of the Acoustical Society of America*, *38*(4), 548–560.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25.

Roberts, L. (1986). Consonance judgements of musical chords by musicians and untrained listeners. *Acta Acustica United with Acustica*, *62*(2), 163–171.

Rodrigues, A. C., Loureiro, M. A., & Caramelli, P. (2013). Long-term musical training may improve different forms of visual attention ability. *Brain and Cognition*, *82*(3), 229–235.

Särkämö, T., Tervaniemi, M., Soinila, S., Autti, T., Silvennoinen, H. M., Laine, M., & Hietanen, M. (2009). Cognitive deficits associated with acquired amusia after stroke: a neuropsychological follow-up study. *Neuropsychologia*, *47*(12), 2642–2651.

Schellenberg, E. G. (2005). Music and cognitive abilities. *Current Directions in Psychological Science*, *14*(6), 317–320.

Schellenberg, E. G., & Trehub, S. E. (1994). Frequency ratios and the perception of tone patterns. *Psychonomic Bulletin & Review*, *1*(2), 191–201.

Schmuckler, M. A. (1989). Expectation in music: Investigation of melodic and harmonic processes. *Music Perception: An Interdisciplinary Journal*, *7*(2), 109–149.

Schoenberg, A. (1975). Style and Idea: selected writings of Arnold Schoenberg, ed. *Leonard Stein, Trans. Leo Black (Berkeley and Los Angeles: University of California Press, 1984)*, *216*.

Seachrist, D. A. (2003). *The Musical World of Halim El-Dabh* (Vol. 1). Kent State University Press.

Spence, C., & Driver, J. (1996). Audiovisual links in endogenous covert spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(4), 1005.

Spence, C., & Driver, J. (1997). Audiovisual links in exogenous covert spatial orienting. *Perception & Psychophysics*, *59*(1), 1–22.

Spence, C., & Driver, J. (2004). *Crossmodal space and crossmodal attention*. Oxford University Press.

Squires, N. K., Squires, K. C., & Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology*, *38*(4), 387–401.

Teder-Sälejärvi, W. A., Russo, F. D., McDonald, J. J., & Hillyard, S. A. (2005). Effects of spatial congruity on audio-visual multimodal integration. *Journal of Cognitive Neuroscience*, *17*(9), 1396–1409.

Thompson, W. F., Schellenberg, E. G., & Letnic, A. K. (2012). Fast and loud background music disrupts reading comprehension. *Psychology of Music*, *40*(6), 700–708.

Vos, P. G., & Troost, J. M. (1989). Ascending and descending melodic intervals: Statistical findings and their perceptual relevance. *Music Perception: An Interdisciplinary Journal*, *6*(4), 383–396.

Vroomen, J., & Gelder, B. de. (2000). Sound enhances visual perception: cross-modal effects of auditory organization on vision. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(5), 1583.

Wallin, N. L., Merker, B., & Brown, S. (2001). *The origins of music*. MIT press.

Zehetleitner, M., Koch, A. I., Goschy, H., & Müller, H. J. (2013). Salience-based selection: Attentional capture by distractors less salient than the target. *PLoS One*, *8*(1), e52595.

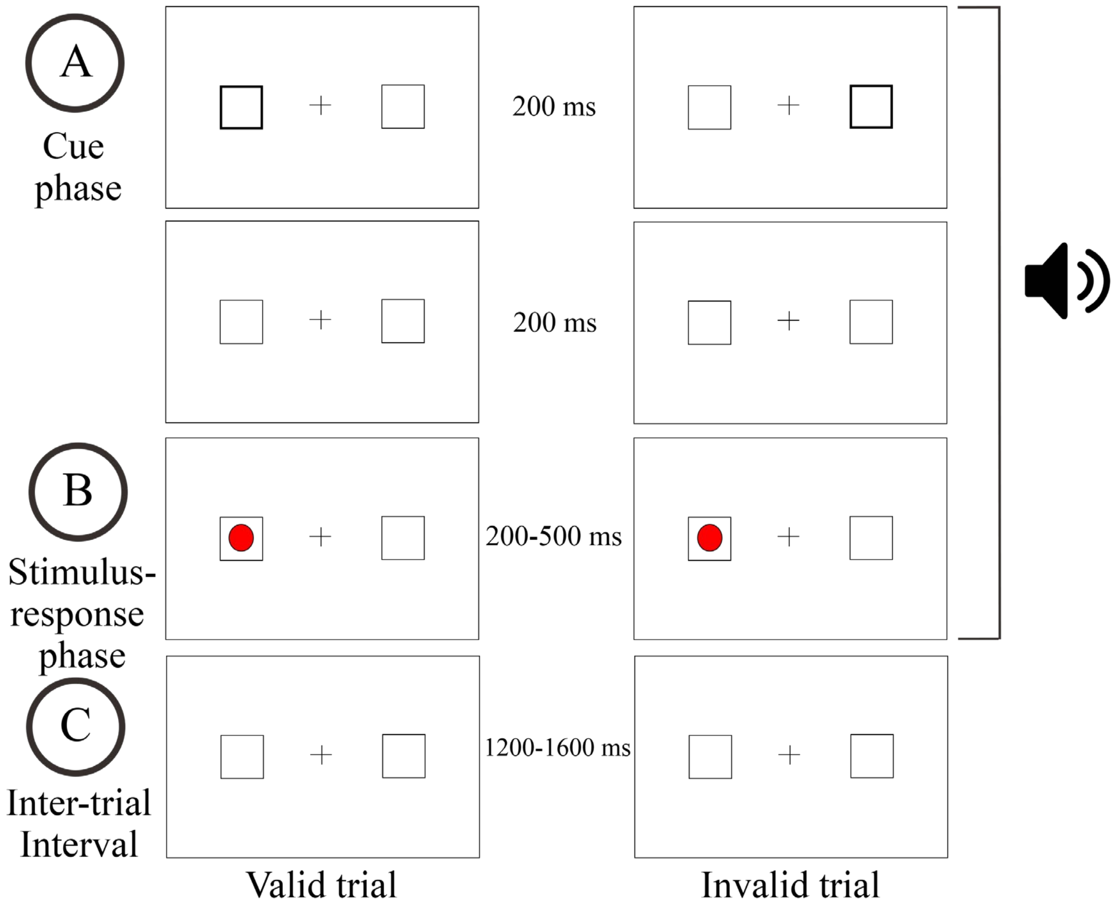

Figure 1. Experimental paradigm. (A) Cue phase, 200 milliseconds (ms). One of two target locations is highlighted, which suggests the appearance of the target. (B) Stimulus response-phase, 200-500 ms. The target appears in one of the target locations, participant responds with keypress corresponding to target location (i.e. left, right). (C) Inter-trial interval, 1200-1600 ms. A short break in between trials.

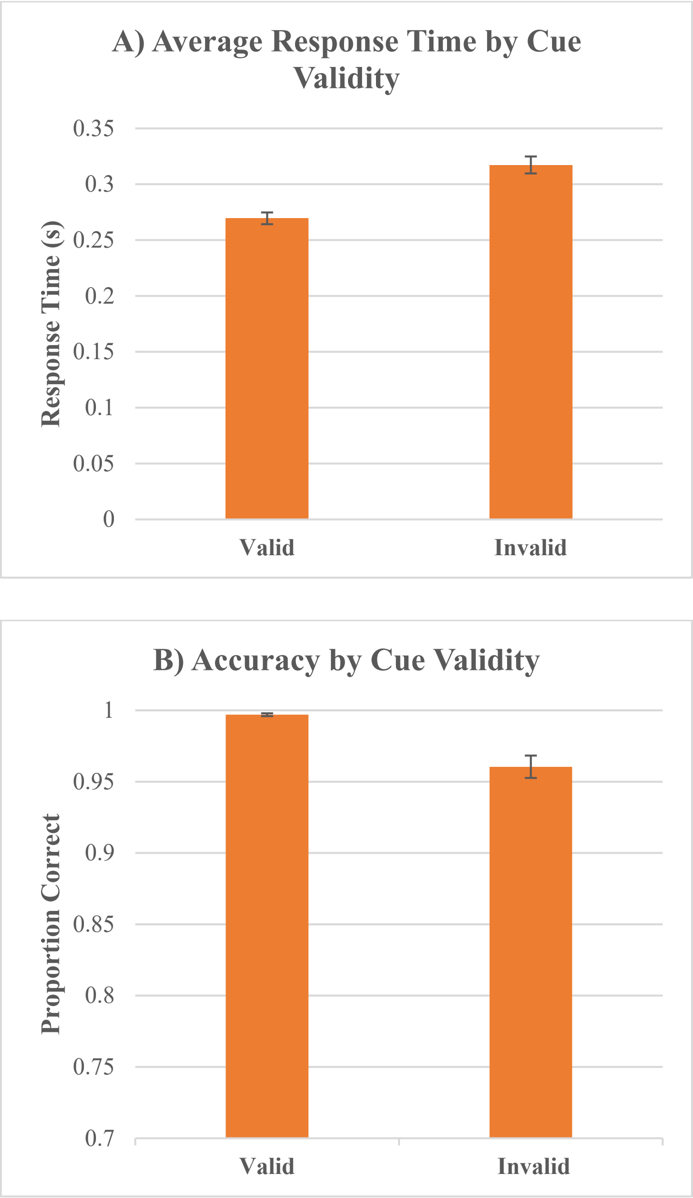

Figure 2. Average behavioral (A) response time and (B) accuracy for valid vs. invalid cue trials.

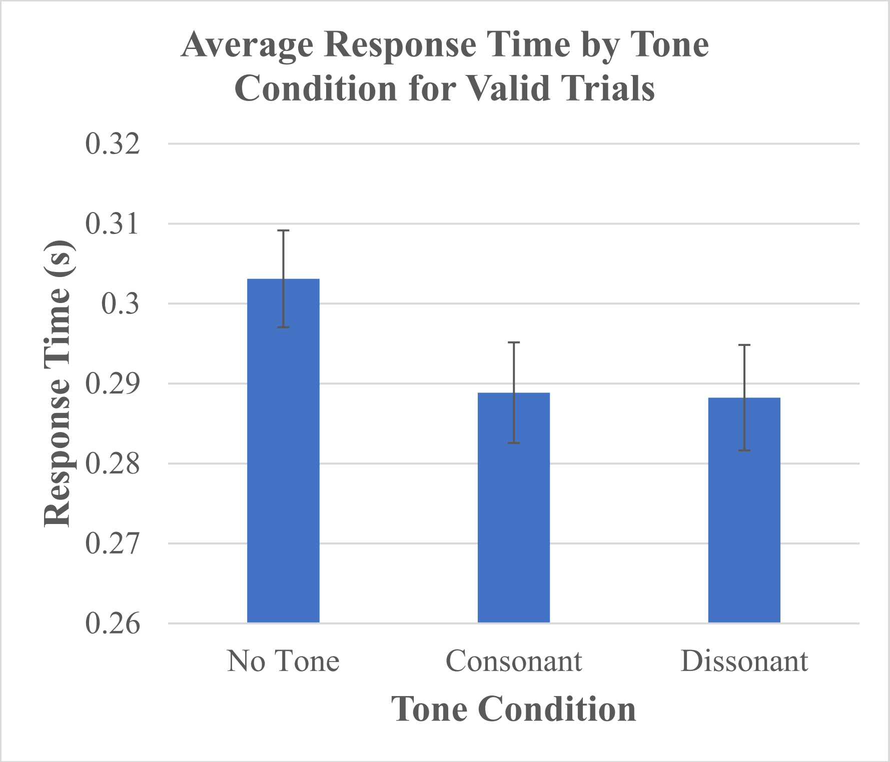

Figure 3. Average response time by tone condition for valid trials.

Appendix A
Program Instructions

In this task, you will be asked to detect the appearance of a target in one of two squares on the left or right side of the screen while keeping your eyes fixed on a central fixation point (+). Please do not look anywhere besides the central fixation point. Before the target appears, one of the boxes may be briefly highlighted. On some trials, a sound may play at the same time that the box is highlighted.

The target will be a RED CIRCLE. If the target appears in the left square, please hit the 'Z' key. If the target appears in the right square, please hit the '/' key. Your goal is to respond as quickly and accurately as possible.

You will now complete a brief practice block. If you have any questions or concerns, please ask the experimenter now. You will also be given an opportunity to ask any additional questions after the practice block.

Appendix B
Administered Questionnaire

Note that this questionnaire was administered after the experimental phase had been completed by the participant.

Which hand do you prefer to use in tasks such as writing?

- ☐ Left
- ☐ Right
- ☐ Neither (ambidextrous)

What gender do you identify with?

- ☐ Male
- ☐ Female
- ☐ Not listed: ___________

Do you have ADHD or any other attentional issues?

- ☐ Yes
- ☐ No

To the best of your knowledge, how many hours of sleep did you get last night? ______

On a scale from 1 to 5, how tired were you before participating this study?

- ☐ 1 - not at all tired
- ☐ 2 - somewhat tired
- ☐ 3 - moderately tired
- ☐ 4 - very tired
- ☐ 5 - extremely tired

Have you played a musical instrument for more than 10 years?

- ☐ Yes
- ☐ No