Detecting Audio-Video Tempo Discrepancies between Conductor and Orchestra

Adam D. Danz (adam.danz@gmail.com)
Central and East European Center for Cognitive Science, New Bulgarian University
21 Montevideo St., 1618 Sofia, Bulgaria

Armina Janyan (ajanyan@cogs.nbu.bg)
Central and East European Center for Cognitive Science, New Bulgarian University
21 Montevideo St., 1618 Sofia, Bulgaria

Abstract

Tempo perception has been studied using many different methodologies in trying to isolate the motion patterns of a conductor in experimental stimuli. In doing so, the context of conducting has been ignored including secondary bodily motion, expression and communication from the conductor, and conducting gestures outside of the realm of tempo control. The aim of this study is to test the ability of trained musicians to detect tempo discrepancies between a conductor and an orchestra using natural footage of a conductor from a musician's perspective under manipulated experimental conditions.

Keywords: tempo discrepancy; conducting; orchestra; perception; audio-visual; RT.

Introduction

The ability of humans to subdivide a flowing continuum of music into equal and regular temporal components known as tempo has been studied using many different experimental paradigms (Lapidaki, 2000; Moelants & McKinney, 2004; Repp & Keller, 2004; Wang, 1984). However, in efforts to isolate independent variables and eliminate confounding variables, experimenters have created stimuli that no longer represent realistic performance conditions. Stimuli have been reduced to mere auditory metronomic pulses, conductor motions have been restricted and confined to the isolated motion of arms or hands, and only small segments of music have been used in stimuli not representing the experience of movements and phrases heard in music performance.

The present study investigates the ability of trained musicians to detect audio-visual tempo discrepancies between a conductor and an orchestra by using natural and unrestrained video footage of a conductor accompanied by an orchestra. The video was then manipulated to create the experimental conditions. This study is not only the first to use a natural setting of conductor and orchestra but also the first to investigate detection of tempo discrepancies via manipulation of visual modality in simultaneous comparison to a steady audio tempo.

Tempo Ranges and Preferred Tempo

Tempo in music is measured by beats per minute (bpm) and represents a division of time that cycles for a finite period during music performance. Just as the human visual system is limited in perceiving only a small segment of the electromagnetic spectrum, the human auditory system is limited in perceiving a certain range of tempi (Van Noorden & Moelants, 1999). A tempo above ~300 bpm (200 ms or 5 Hz) is perceptually difficult to discern due to physical constraints of the human ears and information processing abilities. A tempo under ~40 bpm (1500 ms or 0.67 Hz) results in perceptual isolation of each beat and is beyond the capacity of the working memory to process two consecutive beats that are required to create a tempo. Moelants (2002) implied that there must be a zero point between these ranges where tempo perception is optimal. Using a taping paradigm, Moelants found the range to be centered around 120 bpm (500 ms, 2 Hz) with a range from 115-127 bpm. Using a similar tapping paradigm, Moelants and McKinney (2004) showed that tempi near 120 bpm are more likely to be accurately perceived than tempi further to either extremity. Faster and slower tempi resulted in higher ambiguity of perceived tempo.

Vos et. al. (1997) used a forced directional response paradigm to investigate perceived tempo change in three tempo conditions: (240, 120, and 60 bpm) across 11 incremental conditions ranging from a 10% increase to 10% decrease in tempo. Subjects were presented with auditory stimuli and were instructed to report either an increase or decrease in tempo change. More accelerations were detected than decelerations with the higher tempo value and more decelerations were detected than accelerations in the lower tempo value. An equal amount of accelerations and decelerations were detected with the 120 bpm condition. Vos et. al. (1997) showed that auditory sensitivity to changes in tempo increases as the change in tempo moves further in either direction from 120 bpm.

Similar to the procedure used by Vos et. al. (1997), Jongsma et. al. (2007) performed an ERP study of auditory tempo discrepancies in seven conditions including increase and decrease of tempo discrepancies of 2%, 5%, 10%, and a control condition exhibiting no change. Base tempi were a bit under the 120 bpm 'preferred tempo'. Individual participants differed greatly in their ability to detect small tempo change. Some participants were able to identify all quantities of tempo change above chance level while others were only able to detect the more obvious 10% changes. Interestingly, the direction of tempo change (increase or decrease) had no effect with respect to response accuracy. This might be due to the stimuli falling within a tempo range close to the preferred range described by Moelants. (2002).

Incorporating visual tempo

There are many sources of tempo in musical performance. Auditory sources include sound produced by other instrumentalists in the ensemble as well as base tempi that could be provided by a metronome. Visual source of tempo include the conductor, the motion of other instrumentalists such as stringed instruments moving their bows in synchrony, or tapping of the foot of another instrumentalist across the ensemble. Furthermore, internal representations of perceived, expected, or interpreted tempi may affect a musician's performance. Luck & Sloboda (2007) tested musicians' ability to synchronize with visual cues provided by a conductor using a point-light paradigm where a reflective marker is attached to each of the conductors' right index finger which is recorded by a motion-capturing camera recording the position of the marker in 3D space at 60 frames per second. Two different conductors with different levels of conducting experience were used to produce the videos. Among the independent variables were slow, medium, and fast tempi (60, 90, 120 bpm respectively). Video recordings were made without the use of a metronome. Participants in the study were instructed to tap in time with the point-light motions. The experienced conductors provided more consistent beat patterns compared to the less experienced conductor. Surprisingly, the less experienced conductor elicited slightly more consistent synchronization from participants compared to the experienced conductor. With concern to the tempi, the smallest mean standard deviation (MSD) was associated with the fast tempo (120 bpm) and the largest MSD was associated with the slow tempo (60 bpm).

The point-light paradigm has been often used in music research but lacks contextual data that is pertinent in music performance. Luck & Slaboda's (2007) stimuli, for example, only tracks the right hand leaving the left hand of the conductor, often used for dynamic control, beginning and ending of phrases, and expression (Holden, 2003), unaccounted for. Not only this, but point-light paradigms representing conducting patterns excludes other bodily motion, facial expression, and eye contact used by conductors to communicate and express to the musicians.

Incorporating Context

Whereas most experimental paradigms of tempo discrimination involve isolation of audio or visual tempi from other factors by using metronomic beats or point-light methods, in an attempt to add context into an experimental paradigm, Wang (1984) programmed a symphonic excerpt into a computer so that tempo of a musical composition could be altered and tested. Other factors were tested as well including even and uneven rhythm, solo melody and accompanied melody, etc. Each item began with 72 bpm and remained steady for at least eight beats. 88 college students were asked to tap their fingers while counting the number of beats and to note when the change of tempo occurred and if the change was faster or slower. Tempo changed at +/- 1 bpm per measure after onset. "…subjects needed more time to perceive tempo increase than tempo decrease, more time to perceive tempo change when the rhythm was uneven than even, and when the melody was played alone than when accompanied with steady rhythm." (p. 174). Furthermore, 4% of the responses indicated a tempo change before the experimentally manipulated change actually took place. Similar premature responses occurred in the present study and will be discussed later.

Clemens (2008) investigated the perception of expressive movements used by conductors as seen from three different positions using five different conductors. Participants with orchestral performance experience watched video sequences of a conductor from three different positions: left hand side of the conductor, a frontal view, and to the right of the conductor. Two pianists performed music with the conductor. Participants judged the expressiveness of the conductor with a continuous response interface and then rated several items including amount of information, level of arousal, valence, expressive communication, facial expression, rhythmical clarity, and charisma. Amount of information, level of arousal, and rhythmical clarity resulted in higher ratings from the frontal perspective while expressive communication was rated higher from the left-hand perspective compared to the right-hand perspective. Significant differences also occurred in these dimensions between conductors. There were no differences in the findings according to the participants' usual positions in an orchestra given his or her particular instrument. Therefore habituated perspectives of the conductor had no effect when viewed from other angels.

Considering the above, the present study investigates the interplay of audio and visual tempi as experienced by a musician in an orchestral setting with real instrumentation and a full view of a conductor (from the waist up). Unlike previous studies, audio aspects remained the same while the visual timing was manipulated by an increase and decrease of 5 and 20 percent tempo adjustments. All subjects viewed the conductor from the same perspective at about fifteen degrees to the left of the conductor which, according to Clemens' findings, offers maximal information to the musician of tempo and expression. The change of perspective from the participants' normal placement in an orchestra according to their instrument to the current fixed perspective used in this study should also have no effect.

Method

Participants

Sixteen musicians (7 males) participated in the experiment including 5 violinists, 5 vocalists (three soprano, one tenor, and one alto), 3 pianists, 2 trumpeters, and a trombonist with an overall mean age of 32 years (19-50 years; 9.4 SD). Of these participants, three had previous experience conducting, one of which reported a high level of conducting experience and ability. The only requirement for participation was a minimum of two years experience in an orchestra or choir lead by a conductor. See table 1 for summary of participant data.

Table 1: Characteristics of the participants, means and standard deviations in parenthesis.

	M (SD)
Self estimated proficiency in instrument	5.31 (1.22)
Self estimated proficiency in understanding a conductor	5.56 (1.07)
Years of performance in an orchestra or choir	10.63 (7.00)
Years of experience with perspective instrument or voice	17.88 (9.10)
Age	32 (9.40)

Note. Proficiencies were self-rated on a seven point scale where seven is most proficient.

Stimuli

180 short video clips were segmented from a six-minute video of a graduate student from Texas State University conducting an orchestral version of O Magnum Mysterium arranged by Morten Lauridsen (2000) at about 76 bpm. The original video was compressed with low resolution. The view of the conductor was stable throughout the video at an estimated angle of 15 degrees to the conductor's left with full frontal view of the conductor from waist up. No musicians were in view and the conductor's waist-to-head stretched from bottom to top of the visual display while his arms and hands were in full view throughout each clip. From these six minutes, 180 clips were segmented into 17-33 seconds each. Segments therefore overlapped with other segments but never shared the same starting point nor point of tempo discrepancy. Certain parts of the video were avoided in creating the stimuli including refrains, tempo changes, and whole notes played by the ensemble since they do not exhibit an auditory tempo. Each clip began on the downbeat, or 'beat one', of a measure or within a few milliseconds of the downbeat.

Using Ulead VideoStudio 8.0 SE Basic, the audio portion of each clip was separated from the video. At least four beats into each clip, or four conducting strokes, a downbeat was selected as a point of discrepancy where the video aspect was increased and decreased at five and twenty percent making a total of four clips of the same segment of music in four separate conditions (2x2 (direction of tempo change: increase vs. decrease) x (magnitude of tempo change: 5% vs. 20%)). Each clip ended ten seconds after the point of discrepancy allowing for the same temporal window for detection for each clip. However, depending on the magnitude and direction of the tempo change, this window will vary in the number of visual beats after the discrepancy from 10 (decrease of 20%) to 15 beats (increase of 20%).

The clips were then reformatted into MPEG format suitable for use with the E-prime software (Schneider, Eschman & Zuccolotto, 2002). Of the 180 clips, 10 were practice trials and 12 were left unchanged and without discrepancy. Two of these clips were used as practice trials and as unchanged trials. The remaining 160 clips were the 40 experimental conditions represented in all four conditions. The trials that did not have an experimentally manipulated tempo discrepancy varied in length from 17-31s.

This particular video was chosen because of the moderate experience level and low familiarity of the conductor and musicians. The low resolution of the original video was also desirable in that it completely camouflaged the visual lag effects that usually accompany temporal editing of video.

Procedure

Subjects were tested individually on a Dell laptop computer (1.70 GHz processor with 1 GM RAM) with E-prime software (Schneider, Eschman, & Zuccolotto, 2002). Subjects sat comfortably in front of the laptop wearing full coverage headphones with volume control and an external mouse with one button. Subjects were provided with written instructions to press the mouse button if and when they feel that there is a discrepancy between the conductor and the orchestra. Subjects were informed by the experimenter to listen to the ensemble as a whole rather than concentrating on individual instruments. Additionally, subjects were instructed to answer as quickly as possible but to focus on accuracy and to be sure that they have detected a discrepancy before depressing the button. Four pseudorandomized lists were generated so that no more than two stimuli of the same conditions occurred in concession. Non-manipulated trials were placed throughout the list. Each trial began with a 500ms fixation cross followed by the clip. The intertrial interval was set at 1500 ms. After 10 practice trials, subjects read a second instruction to ask questions if there is anything unclear. The experimental session then followed lasting between 20-28 minutes.

Results

Dependent variables included reaction time (RT) and percentage of error in detecting the discrepancy between visual and audio tempi. From the 800 total trials, 160 were non-manipulated control trials. The remaining 640 were experimentally manipulated independent variables (320 in each factor of increase and decrease). Of these 640 tempo discrepancies, 27.5% were missed by participants and no RT was recorded. From the remaining 464 trials that exhibited detection, 317 (49.5% of the whole number of trials) resulted in detection after the onset of the stimuli while the remaining 147 trials (23%) exhibited detection before the onset of the stimuli. These data suggest that either 1) the participants were not good enough in detecting audio-visual tempo discrepancy, 2) the stimuli exhibited unintentional tempo discrepancies by the musicians and/or conductor, 3) quality of the video effected participants' judgment, or 4) any combination of these factors.

Reaction time analysis was based on correct responses only. For a response to be correct, discrepancy detection was required to occur within the ten-second window after the discrepancy onset. This eliminated trials that did not feature a discrepancy as well as trials where participants clicked the button before the discrepancy onset. 23% of the trials resulted in a detection before the onset of the manipulated discrepancy. There are several explanations for this relatively high error rate. First, as discussed previously, the quality of performance of the conductor and the orchestra was moderate at best. The beginning of some phrases were not directly on cue with the downbeat of the conductor. The performance of the brass section was a bit weak at times as well. Also, the conductor naturally exhibited slight tempo variances from time to time, typical of live performance. Therefore the participants could have had valid reason to press the button in detecting the smaller fluctuations of tempo that were outside of experimental control. This may be a result of using contextual stimuli extracted from real performance rather than experimentally controlled or simplified stimuli.

Another cause of possible tempo ambiguity is the low resolution of the video which made the conductor's larger arm movements appear to have a slight lag compared to the smaller detailed movement. Lastly, due to being instructed to detect tempo variations, subjects put extra attention into tempo and may have been sensitive to any type of fluctuation in tempo whatsoever. For these reasons, detection of a discrepancy before the experimentally manipulated onset was not counted as error nor measured in reaction time analysis.

Table 2: Mean response times (in ms) and standard deviations in parentheses for four experimental conditions, item means.

	5%	20%
increase	6994(1871)	4420(1760)
decrease	6396(3409)	4922(1276)

Table 2 shows mean reaction time and standard deviation per condition. Standard deviation scores are high due to high variability between participants and their formal experience and training. Conducting gestures are highly individual requiring much exposure to a single conductor before musicians are able to read and understand his non verbal communication. This particular conductor was unknown by any of the participants who were reading his gestures for the first time. Secondly, the musical education and experience of the participants varied greatly. There were five participants acquainted with string instruments, three with brass, three piano, and five vocalists. 31% of the participants therefore have most experience with choir conducting which varies in style from orchestral conducting. Though O Magnum Mysterium is originally a choir piece (Lauridsen, 2000), it was performed only with brass instruments in this study and no participant had recognized or previously performed the piece.

A 2x2 ANOVA analysis on item and subject RT means obtained main affect of magnitude of change (20% vs. 5% change) (F_i (1, 134)=31.83; $p<0.001$; F_s (1, 56)=13.31; $p<0.001$) which suggests that participants detected 20% tempo changes faster (4665 ms) than 5% changes (6754 ms). The first reason for high RT is because of the somewhat slow base tempo of 76 bpm of the stimuli. Two consecutive beats at this tempo consume 790 ms which would be the minimal time necessary to identify the 76 bpm tempo. Secondly, unlike other tempo detection experiments, subjects were instructed to focus on a visual tempo change in comparison to an audio tempo counterpart *being perceived simultaneously*. This requires a vast amount of working memory and timing processes that are not necessary in single-tempo processing. Lastly, the high RT values may suggest that the process of detecting visual tempo-changes at this tempo is neither simple nor automatic and may require additional processes or mechanisms to confirm the change in tempo. No main effects of directional tempo change (increase vs. decrease) or interaction were obtained in either item or subject analysis ($p>0.1$). In other words, participants exhibited quicker reaction time to 20% visual tempo changes than 5% changes but the direction of the change (increase or decrease) had no effect. (cf. Table 2).

Table 3: Mean error rates (in %) and standard deviations (in parentheses) for four experimental conditions, item means.

	5%	20%
increase	36 (33)	18 (19)
decrease	73 (29)	13 (22)

Error analysis only included experimentally manipulated stimuli that were left undetected by participants. Trials where subjects pressed the button before the programmed discrepancy were not included as error, for the subject may have detected the smaller unintentional perturbations of tempo by the conductor. A 2x2 ANOVA analysis on item and subject error means obtained both main effect and interaction. A main effect of direction of tempo change ($F_i(1, 155)=14.69$; $p<0.001$; $F_s(1, 60)=5.79$; $p<0.05$) showed that participants made less errors in the tempo increase condition (27%) than in the decrease condition (44%). A main effect of magnitude of change ($F_i(1, 155)=85.42$; $p<0.001$; $F_s(1, 60)=38.49$; $p<0.001$) replicated the RT effect showing that participants made considerably fewer errors in the 20% change condition (16%) than in the 5% condition (54%). Finally, an interaction between direction of tempo change and magnitude of change ($F_i(1, 155)=25.00$; $p<0.001$; $F_s(1, 60)=11.64$; $p<0.01$) confirmed the strong magnitude of change effect and showed that the tempo change effect is present only in the 5% change condition (cf. Table 3 for error means). Thus, the most difficult condition for participants was the 5% decrease of tempo change. The least difficult, predictably, was the 20% change in magnitude of discrepancy. The results unambiguously suggest that 20% change in tempo overrides the directional manipulation of increase and decrease of tempo.

Discussion

The present study differs from other tempo discrepancy research in that audio tempo remains stable while visual tempo cues are what is experimentally manipulated. Furthermore, subjects were required to simultaneously process more than one tempo via two different modalities. Therefore, results should supplement previous research and care should be taken in comparing these results with experiments that measure audio tempo perception alone. Furthermore, stimuli used for the present study were extracted from an actual orchestral performance not arranged for experimental purposes. While careful consideration was taken in manipulating and controlling the independent variables, many confounding variables existed in order to portray a contextually accurate orchestral setting.

The present study did not concur with the findings of Vos et. al. (1997) as well as Moelants (2002). In the present study, participants made fewer errors in the tempo increase condition (27%) than in the decrease condition (44%). Being that the base tempo for all stimuli in the present study was around 76 bpm, detection of tempo *decrease* should have had higher accuracy and perhaps quicker RT according to Vos et. al. (1997) because a decrease in tempo from 76 bpm is going farther from the 'preferred tempo' and therefore should be more easy to detect. Furthermore, in Vos et. al. (1997), audio tempo discrepancies diverging from 60 bpm resulted in higher accuracy for decelerations of tempo than accelerations. Only 16 bmp greater, the present study resulted in higher accuracy of detecting increases of tempo discrepancy. This may be because Vos et. al. (1997) greatest magnitude of tempo change was 10% whereas the present study manipulated tempo at 20%. Furthermore, it can not be emphasized enough that a comparison is being made between audio tempo experimentation (Vos et. al. 1997) and visual tempo experimentation with simultaneous audio tempo perception (present study).

Participants in the present study made considerably fewer errors in the 20% change condition (16% error) than in the 5% condition (54% error). A 20% change from 76 bpm results in an instant jump to either 61 or 91 bpm which is considerable faster or slower from the original tempo. A 5% change results in only +/- ~4 bpm which is naturally more difficult to detect (Vos et. al., 1997). 20% temporal changes made during the editing process will naturally affect the quality of video but once again, the low resolution of the original video camouflaged these effects and no subject reported detection of video-editing artifacts.

The current results support the findings of Moelants & McKinney (2004) in that ambiguous tempi are more accurately perceived as the base tempi approach 120 bpm while tempi at further extremities results in higher ambiguity. Tempo increases from the base tempo of ~76 bpm resulted in less error. Also noted in Moelants & McKinney (2004) and relevant to the present study, structural characteristics of music such as deviating accents, etc, can add to ambiguity of perceived tempo.

As Wang (1984) has demonstrated, tempo is not the only factor that effects processing time of tempo discrepancies. Even and uneven rhythm affected RT in detection of tempo change as well as solo vs. accompanied melody. While rhythm remained even for all stimuli used in the current study, melody and instrumentation varied as some segments featured solo horn sections and others featured accompanied melody. Detection of tempo discrepancies before the onset of the experimentally manipulated discrepancy was common in both the current study (23% extraneous detections) and in Wang (1984) (4% extraneous detections). The higher extraneous discrepancy detection in the current study may have resulted from the previously discussed confounding variables (video quality, skill level of the conductor, performance of the musicians, etc.). Additionally, Wang used programmed melodies played by a computer without visual tempo (i.e. conductor) whereas the present study used contextual video of a conductor and orchestra requiring dual concentration across two modalities and processing and comparison of two tempi in each trial.

Another source of difficulty for participants could have been the slow tempo of each stimulus. Luck and Slaboda (2007) had already confirmed that synchronization with slower tempi (60 bpm) resulted in less consistency with subjects who were instructed to tap along with the tempo. Being that the current study is within this slower range of tempi, perhaps synchronization with the auditory aspect (the orchestra) was difficult making the task of following the visual tempo (the conductor) even more challenging. Whereas the previous studies have isolated one type of tempo (audio or visual), the present study requires participants to put forth attention to both modalities without specific instruction to concentrate on a particular modality. This causes an increase in processing and, along with the slower tempo range, an increase in working memory.

While the quality of video used in the stimuli was not of high resolution, the conductor's arms and motions were always in view and completely discernable. Furthermore, along with Clemens' (2008) findings, the angle of view of the conductor provided maximal information for both tempo and expression to the viewers. What is more, Clemens' findings state that habitual viewpoint of the musicians (i.e., where they usually sit in the orchestral setting) should have no effect on the fixed viewpoint in the stimuli.

The educational levels of the conductor in the current study and the conductor used in Luck & Slaboda's (2007) study both are at a master class level of performance at the time of recording. Luck and Slaboda found that the less experienced conductor aforementioned elicited more consistency in synchrony from subjects compared to the more experienced conductor they used in their study. This provides support that the lower experience level of the conductor in this current study should not affect participants' synchrony of tempo.

Lastly, as in Jongsma et. al. (2007), individuals differed greatly in detection of tempo changes. 31% of participants in the present study were choir members who are typically exposed to a slightly different style of conducting than orchestral musicians. This is another advantage of the lower experience level of the conductor used as he has not yet adopted the strong characteristics of orchestral conducting and his gestures can be interpreted in varying genres.

The data does not concur with previous studies and many of the trials were abandoned in statistical analysis due to extraneous tempo discrepancy detections. This is most likely a result of the stimuli used (i.e. the quality of recording, efficacy of conductor, and abilities of musicians performing in the orchestra) but may be a result of visual tempo detection ability. Simplified stimuli should be created for future studies but without the elimination of contextual attributes. Higher resolution would be desirable as would a more proficient conductor and orchestra. Furthermore, audio-visual tempo discrepancy should be investigated holistically whereas the present study only manipulated visual aspects while auditory tempo remained constant.

Acknowledgments

This research was supported financially by the ANALOGY project (NEST program, contract 29088) funded by the EC. We thank our participants for their patience, Svetoslava Georgieva for use of equipment and three anonymous reviewers for their useful comments and suggestions.

References

Clemens W., C., Auhagen, W. (2008). Perceiving Conductors' Expressive Gestures from Different Visual Perspectives. An Exploratory Continuous Response Study *Music Perception 26* (2), 129–143.

Holden, R. (2003). The technique of conducting. In J. A. Bowen (Ed.), *The Cambridge companion to conducting*. Cambridge: University Press.

Jongsma, M.L.A., Meeuwissenc, E., Vosd, P.G., Maes, R. (2007). Rhythm perception: Speeding up or slowing down affects different subcomponents of the ERP P3 complex. *Biological Psychology 75* (3), 219-228

Lapidaki, E. (2000). Stability of tempo judgments in music listening. *Music Education Research*, 2(1), 25-44.

Lauridsen, M.J. (2000). O Magnum Mysterium [Score]. Southern Music Publishing Co. Inc.

Luck, G., Sloboda, J. (2007). Synchronizing with complex biological motion: An investigation of musicians' synchronization with traditional conducting beat patterns. *Music Performance Research*, *1*(1), 26–46.

Moelants, D. (2002). Preferred tempo reconsidered. *Proceedings of the 7th International Conference on Music Perception and Cognition*, Sydney

Moelants, D., & McKinney, MF. (2004). Tempo perception and musical content - what makes a piece fast, slow or temporally ambiguous. *8th intern. conference on music perception & cognition*, Evanston, IL.

Repp, B. H., & Keller, P. E. (2004). Adaptation to tempo changes in sensorimotor synchronization: Effects of intention, attention, and awareness. *Quarterly Journal of Experimental Psychology*, 57A, 499-521.

Schneider, W., Eschman, A., Zuccolotto, A. (2002). *E-Prime user's guide.* Pittsburgh: Psychology Software Tools Inc.

Van Noorden, L., Moelants, D. (1999). Resonance in the perception of musical pulse. *Journal of New Music Research*, *28*(1): 43-66.

Vos, P. G., van Assen, M., Franek, M. (1997). Perceived tempo change is dependent on base tempo and direction of change: Evidence for a generalized version of Schulze's (1978) internal beat model. *Psychological Research 59*, 240- 247.

Wang, C. (1984). Effects of some aspects of rhythm on tempo perception. *Journal of Research in Music Education*, 32, 169-176.