Low Motion Smoothness of Dance leads to Slower Detection of Auditory Target

Ahhyun Lee (a), Jinhee Kim (b), YouJeong Hong (c), Kyung Myun Lee (d), Kyogu Lee (e)

(a) Pohang university of Science and Technology, South Korea, ahhyun724@postech.ac.kr, (b) Seoul National University, South Korea, ginnykim9@snu.ac.kr, (c) Seoul National University, South Korea, yjyjyj12@snu.ac.kr, (d) Korea Advanced Institute of Science & Technology, kmlee2@kaist.ac.kr (e) Seoul National University, South Korea, kglee@snu.ac.kr

Dance accompanies musical rhythm along with motion. Previous research has explored the relationship between auditory rhythm processing and the visual movement. Lee *et al.* (2015) showed that participants' response times to auditory deviant occurring at strong metrical positions were significantly slower when viewing dance compared to an animated picture. This implies that dance can hinder simultaneous rhythm processing by directing attention towards the visual modality. The current study focused on a specific visual attribute of dance, motion smoothness (MS), the fluidity of motion changes. Response times to deviant tones presented within regular or irregular rhythm sequences accompanied by dance movements with high or low MS were measured. The results showed that response times were longer in the low MS condition compared to the high MS condition, while rhythm regularity had no significant effects. These findings suggest that abrupt changes in dance motion speed (low MS) may impose an additional perceptual load, causing individuals to allocate more attention to the visual modality and, consequently, to exhibit slower responses to concurrent auditory targets.

Keywords: Dance Perception, Motion Smoothness, Auditory Processing, Divided Attention, Visual Distraction, Audiovisual Perception

1. Introduction

Music has the power to induce movements, from tapping to dance movements. As Leman (2007) suggests, such spontaneous movements might arise from human's ability to predict the upcoming beat and rhythm (Burger *et al.*, 2014). The ability to entrain to the rhythmical structure (i.e., rhythmic entrainment) has been known to be a unique feature of humans and certain groups of bird species (Merchant & Honing, 2014). Considering the situation of viewing dance, as an audiovisual stimulus, the visual dance movement and concurrent musical rhythm are conveyed through audiovisual integration process.

Not only rhythm and dance, but there have also been numerous studies demonstrating that concurrent visual information can influence sound perception. For instance, studies investigating alterations in sound perception during music performance have shown that the factors, such as, note duration (Schutz & Lipscomb, 2007), the size of sung intervals (Thompson *et al.*, 2010), and quality of music performance (Tsay, 2014), can be perceived differently when accompanied by performers' body movements. Considering that changes in visual stimuli can alter perception of simultaneously presented sound, it is plausible that elements of dance stimuli can lead to altered perception to visual and auditory modality.

Previous research has demonstrated a relationship between the perception of auditory rhythm in dance videos and dance movements. Several studies have shown that dance movements can induce auditory beat perception, such as street-dance-like bouncing movements facilitating auditory beat perception (Su, 2014a), and swing dance eliciting a sense of metrical structure in visual rhythm (Su, 2016). Conversely, other research has shown the ineffectiveness of visual information in beat induction. People have exhibited poorer synchronization to visual metronome compared to auditory metronome (Patel *et al.*, 2005; Repp & Penel, 2002), in beat perception (Grahn *et al.*, 2011) and rhythmic interval timing tasks (Grondin & McAuley, 2009). However, improvements in synchronization and rhythm perception were observed when spatiotemporal information was conveyed through visual stimuli, such as moving bars and bouncing balls. (Hove *et al.*, 2013; Hove *et al.*, 2013) These findings highlight the significance of spatiotemporal information in visual beat induction. The present study aims to focus on a specific attribute of the spatiotemporal factors in dance, motion smoothness, which we hypothesize to be important for visual beat induction.

Lee *et al.* (2015) demonstrated that the difference in response time (RT) to auditory deviants along metrical position, representing a hierarchical organization of beats, was more evident in dance stimuli compared to picture condition. This suggests that dance stimuli have a greater influence on directing attention towards a stronger metrical position. The non-significant effect observed in picture condition is attributed to the lack of spatiotemporal information in the visual stimuli, which in turn has a limited impact on simultaneous sound perception. Other evidence supports the notion that temporal information, such as acceleration rate, plays a crucial role in visual beat production. Luck & Sloboda (2009) found that visual beat induction from conducting gestures is primarily related to acceleration along the trajectory. Thus, the present research aims to examine the impact of a specific spatiotemporal element within dance movements on auditory perception. Specifically, we propose that *motion smoothness* of dance and audiovisual synchrony are potential factors that influence auditory rhythm perception by affecting visual saliency. We define motion smoothness (hereafter, MS) as the extent to which motion speed changes fluidly (high MS) or abruptly with sudden changes (low MS). Orlandi *et al.* (2020) demonstrated that different movement kinematics within the same dance movements can lead to varied aesthetic appreciation. Two types of dance kinematics were examined: *uniform* kinematics, characterized by consistent movement speed throughout the entire sequence, and *varied* kinematics, emphasizing dynamic changes in movement speed. Since MS can be measured as the number of velocity peaks per meter in movements with equal trajectories (Balasubramanian *et al.*, 2015), uniform kinematics can be described as dance movement with high MS, while varied kinematics can be described as dance movement with low MS, resulting in salient moments of acceleration and pause.

In addition to exploring MS in visual dance, we also explored audiovisual synchrony as a potentially crucial factor in inducing audiovisual beats. When the onset of auditory and visual stimuli does not coincide and its stimulus onset asynchrony exceeds the temporal binding window, it is less likely for the stimuli to be perceived as a unified event (Bidelman, 2016; Chen & Spence, 2017). Therefore, the asynchrony between visual dance and auditory rhythm could influence the processing of audiovisual rhythm. Research has shown that the synchrony between dancer movements and accompanying sounds enhances the detection of point-light tap dancers in the presence of visual or auditory noise (Arrighi *et al.*, 2009). However, in tasks with lower attentional demands, the synchronous condition did not yield a significant benefit (Sevdalis & Keller, 2011). Since the audiovisual synchrony determines whether audiovisual stimuli are perceived as a unified source, it represents an important factor in conjunction with individual characteristics of auditory stimuli and visual stimuli.

Our research aimed to investigate the potential impact of MS in dance and audiovisual synchrony between dance and auditory rhythm on RTs for auditory deviants in concurrent auditory rhythm. We created videos of movements with the same motion trajectory but different MS to manipulate their saliency over time. Participants were instructed to press a button upon the auditory deviant within base rhythm presented with visual stimuli, which reflects their auditory attention. Based on previous literature, two hypotheses can be proposed. If visual information facilitates auditory perception, visual stimuli with low MS would lead to increased attention towards auditory deviants, resulting in shorter RT for low MS condition. Su (2014b) demonstrated that bouncing human point light figures conveying visual beat information enhanced the ability to perceive and synchronize to auditory rhythms. According to the dynamic attending theory, when individuals can anticipate or predict upcoming temporal events, they allocate attentional resources to the predicted time points, leading to more efficient processing (Bolger *et al.*, 2013). In this context, visual stimuli with low MS would elicit time points with higher expectation for salient beats, resulting in faster responses to auditory deviants in salient visual beats. On the other hand, if the limited attentional resources are temporally divided between auditory and visual domains, visual stimuli with higher MS might result in longer RTs, as more attentional resources are allocated to processing the visual domain. It has been suggested that whether each sensory modality has its own attentional resources, or whether different sensory modalities share attentional resources, depends on the task. The majority of studies suggest that attentional resources for auditory and visual modalities are different, and that timd-limited object-based attention tasks require the recruitment of shared attentional resources.(Wahn & König, 2017). Lee *et al.* (2015) showed that slower RTs in the dance condition compared to the picture or control condition may be attributed to the presence of the dance video, which increased perceptual load. Although the two perspectives support different mechanisms of audiovisual integration, they both support that visual stimuli with low MS would have higher visual salience, either enhancing or hindering auditory processing.

2. Methods

2.1.Participants

Thirty participants (13 males; M_{age} = 29.433, SD_{age} = 2.04 years) with normal vision and normal hearing were recruited. Prior to participating in the experiment, all individuals provided written informed consent. The research obtained approval from the ethics commission of the Seoul National University Institutional Review Board.

2.2. Materials

The visual stimuli were recorded against a white background without any additional objects. The visual stimuli consisted of 2 sets of movements that varied in terms of motion smoothness, while sharing the same average angular speed and motion trajectory. There was a total of 6 types of visual stimuli, combining 3 dance motions (upper torso, chest, forearm) with 2 levels of motion smoothness (high, low). The dance motions corresponded to specific body parts being moved. The first type of dance involved repetitive movements of the upper torso in four directions: front, right, back, left, while maintaining a stationary position. The second type of dance featured the repetitive movement of the chest in four diagonal directions: front-left, back-left, back-

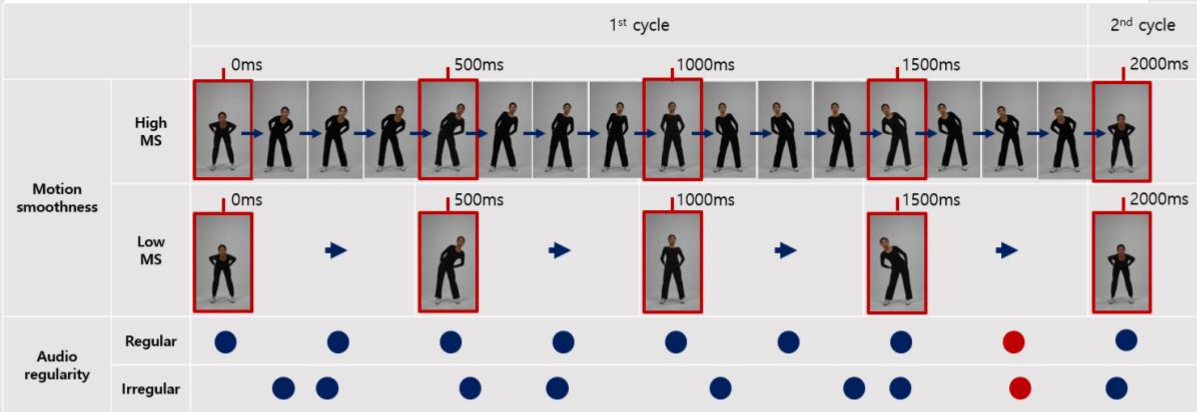

Figure 1. Schematic description of experiment: Dance motion moving upper torso is described with audio regularity conditions. Dark blue dots indicate base rhythm and red dots indicate auditory deviant. Stimuli archive can be accessed by the link below.
https://bit.ly/42eXTvp

right, and front-right. The third type of dance comprised the repetitive movements of the forearm in four directions: up, left, low, and right. Furthermore, all visual stimuli were presented at a rate of 120 beats per minute (here after, bpm), repeating every eight beats (2 seconds). These types of dances sharing an identical motion trajectory were categorized into two types based on motion smoothness (hereafter, MS): high MS condition, where actions maintained a constant speed, and low MS condition, where actions varied in speed to emphasize the beat. The dances with lower MS exhibited salient visual beats through sudden changes in moving, while dances with higher MS displayed less discernible visual beats with a constant moving speed.

The auditory stimuli consisted of a base rhythm designed to evoke rhythmicity, accompanied by a target tone. The base rhythm was presented in a different manner depending on the audio regularity condition. In the regular condition, an isochronous rhythm with an IOI of 250ms (corresponding to 120 bpm) was synchronized with the visual stimuli, which was repeated every 2 seconds. In the irregular condition, the base rhythm had an IOI randomly selected from a normal distribution (mean = 50ms, SD = 25ms). The base rhythm generated by Audacity software consisted of a sound with a timbre of "tick" and a pitch of D3 (146.83Hz). The target tone with a timbre of "beep" was presented intermittently along with the base rhythm and had a pitch of F#6 (1479.98Hz). Each trial lasted for 100 seconds and comprised 400 presentations of the base rhythm and 28 instances of the target tone. The target tone was simultaneously presented with the base rhythm at randomly selected positions generated by python, according to the following rules. The target tone only appeared at the first, third, fifth, or seventh position within the 8-beat base rhythm cycle, which lasted 2 seconds. An IOI of the target tone was set to be at least 2 seconds, and the initial 2 seconds of each trial did not contain any target tones to allow participants to get used to the stimuli.

The visual stimuli were filmed at 80 bpm but were sped up to 120 bpm to synchronize with the auditory stimuli, resulting in audiovisual stimuli at 120 bpm. This adjustment was made to ensure the accuracy of the motion and maintain a consistent level of motion magnitude. The first 4 seconds of the audiovisual stimuli consisted of visual stimuli only, allowing participants to familiarize themselves with the visual stimuli itself and the rhythmicity of the visual stimuli, if generated. When visual beat is salient, audio regularity condition can also be considered as audiovisual synchrony condition, while the irregular condition can be considered as audiovisual non-synchronous condition. Audio-only trials were added as a baseline to assess the auditory processing ability in the absence of visual stimuli.

The questionnaires utilized in this study included the Goldsmiths Dance Sophistication Index (Gold-DSI; Rose *et al.*, 2020) and 3 questions assessing music experience. The Gold-DSI was developed to measure individuals' experiences in dance participation and dance observation, which are assessed through participatory score and observatory score, respectively. The participatory score assessed the extent to which individuals engage in dancing (20 questions, min = -28, max = 92), while the observatory score assessed their level of knowledge and viewership of dance (6 questions, score range [-10, 26]). The participatory questionnaire comprised four factors: body awareness, social dancing, urge to dance, and dance training. Additionally, three questions assessing music experience were adapted from the factor 'Dance Training' of the Gold-DSI. These questions included the number of years of daily music practice (M1), the level of experience in music (M2), and the number of years of music education received (M3). (min = 1, max = 7)

2.3. Procedure

Participants underwent a behavioral experiment that entailed the presentation of video stimuli featuring visual dance and auditory rhythms. After the behavioral experiment, they were required to complete the questionnaire.

Participants were asked to press a key for the target tone (i.e., deviant) as quickly and accurately as possible during the presentation of the dance videos and RTs were measured. During a 100-second sequence in each trial, the target tone was randomly presented. Participants were asked to focus on both the visual and auditory stimuli. With a within-subject design, all participants performed total of 14 trials, which included audio only (Ao) trials with two audio regularity conditions (regular, irregular) and twelve audiovisual (AV) trials with three types of dance (upper torso, chest, arm) x two levels of motion smoothness (high, low) and two audio regularity conditions (regular, irregular). Ao trials were presented in a specific order, the regular condition and followed by the irregular condition, to present condition with lower difficulty first. AV trials were presented in random order. Half of the participants began with Ao condition, and the other half began with AV condition. The whole session, consisting of 14 trials of 100 seconds and resting time, lasted around 25 minutes to 30 minutes. After completing the behavioral experiment, participants were required to answer a questionnaire based on the Gold-DSI assessing experience and participation in dance and music.

3. Results

RTs below 100ms and above 500ms were excluded from the analysis to minimize the effect of outliers. The RTs were log-transformed since the overall distribution was right-skewed. In AV trials, when the RTs were averaged across audio regularity conditions, the average RT of the low MS condition was slightly slower than that of the high MS condition. (M_{highMS} = 279ms, M_{lowMS} = 283ms; $M_{highMS\text{-log transformed}}$ = 5.609, $M_{low\ MS\text{-log transformed}}$ = 5.625). A 2 (MS) X 2 (audio regularity) repeated-measures ANOVA was performed on the RTs. The results demonstrated a significant main effect of MS (F(1,29) = 4.525, p = 0.042; M_{highMS} = 5.610, M_{lowMS} = 5.626), indicating that the RTs was longer in the low MS condition. Audio regularity had no significant effects on the RTs both in Ao and AV trials (Ao : F(1,29) = 0.850, p = 0.364; AV : F(1,29) = 1.179, p = 0.286), and the interaction between MS and audio regularity was also non-significant (F(1,29) = 0.131, p = 0.720). Additionally, 2 (AV/Ao condition) X 2 (audio regularity) repeated-measures ANOVA performed on the RTs demonstrated a significant main effect of AV/Ao condition, indicating that the RTs was slower in AV condition. (F(1,29) = 28.202, p < 0.001; M_{Ao} = 5.543, M_{AV} = 5.618)

Table 1
Average RT for MS conditions

	High MS	Low MS
Before log transformed (ms)	279	283
Log transformed	5.609	5.625

A regression analysis between the position of the deviants within each trial and the RT showed that as the deviant position within each trial increased, the corresponding RT decreased, indicating adaptation to task within one trial lasting 100 seconds. This pattern was observed in both Ao trials (R^2 = .20, F(1, 1651) = 426.4, p < 0.001) and AV trials (R^2 = .15, F(1, 9832) = 1827, p < 0.001).

Participants provided information about their experiences of dance and dance observation using an adapted version of the Gold-DSI. Additionally, participants indicated their years of daily music practice, level of experience in music, and years of receiving music education. The average participatory score was 19.56 (median = 18.00, SD = 14.79), and an average observatory score was 7.2 (median = 7.27, SD = 7.475). The average participatory score is below median of the score range and the average observatory score is over median of score range. The average duration of music practice reported by participants was 3.83 years (SD_{years} = 3.94), with five participants reporting a music practice experience exceeding 10 years. The average level of experience in music was between beginner and intermediate. The average duration of music education was 4.13 years (SD_{years} = 3.66), with five participants having more than 10years of music education.

In terms of the questionnaire scores, there was no significant difference. Participants were divided into two groups based on the median of the participatory, observatory and M1, M2, M3 score. A 2 (P score) X 2 (MS) X 2 (audio regularity) repeated-measures ANOVA was performed on the RTs. For all score measures, interaction between score measure and MS was not statistically significant.

4. Discussion

The present study aimed to assess auditory processing efficiency by examining the RTs to auditory deviant tones when presented with visual stimuli. The results of the study revealed a significant effect of MS on RT for the deviant tones, indicating auditory processing efficiency differed in the high MS and the low MS conditions, with slower RT observed in the low MS condition. There was no significant effect of auditory regularity or an interaction between auditory regularity and MS on the RT. The average RT was shorter in the Ao trials than in the AV trials. Furthermore, within one trial which lasted 100-seconds, there was a decrease in RT over time.

The study contributed to advance in theoretical understanding of intertwined patterns of audiovisual processing when people watch dance stimuli with concurrent music. Specifically, it adds to the literature on how processing of auditory rhythm cues can be integrated with concurrent visual cues. We suggested two competing hypotheses: visual enhancement of auditory rhythm and limited attentional resources divided between the auditory and visual modalities. The results partially support the hypothesis grounded in the latter perspective. Within audio regularity condition, longer RTs were observed for the low MS condition, which features salient visual beat compared to the high MS condition. Visual stimuli with low MS would require more attention to the visual domain, resulting in the allocation of fewer attentional resources to the auditory domain and subsequently reducing the efficiency of auditory processing. The present study suggests that dance with lower (vs. higher) MS can increase a perceptual load (Lee *et al.*, 2015), leading participants to allocate more attention to the visual modality and subsequently less attention to auditory modality, resulting in slower responses to concurrent auditory targets. Additionally, the comparison between Ao trials and AV trials also supports this hypothesis. The average RT in Ao trials was shorter than average RT in AV trials, indicating that visual stimuli might have required attentional resources in the visual domain in AV trials compared to Ao trials, thereby reducing auditory processing efficiency.

Delving into a less-explored specific factor in visual dance movement, MS, this study focused on determining which specific factor is important in determining the visual saliency of the dance. Based on previous research that the highlights the importance of spatiotemporal cues in visual beat induction, especially temporal cues in diverse visual stimuli ranging from bouncing balls to conductors' gestures, this study investigated possible temporal cues in dance movement.

Contrary to expectations, auditory regularity and its interaction did not show significance. This may be due to lack of ecological validity or the difficulty of the test, as only "tick" and "beep" sounds were used in this study. Future studies could consider the external validity of the auditory stimuli by using more music-like rhythmic sounds produced by various instruments.

When the results of participants were split by the Gold-DSI and the questionnaire regarding music experience and expertise, none of the responses to the questionnaire showed significant differences. This may be due to mixed effects of dance and music experiences influencing audiovisual integration ability, and the task relying more on bottom-up processing rather than top-down processing. Given that individual differences in dance and music experience can influence audiovisual integration ability (Bidelman, 2016), future investigations can examine the observed effects by recruiting various groups with a range of dance or music expertise.

Considering the observed decrease in RT over time within one trial, this may imply that the task was easy enough for adaptation to occur within 100 seconds. It is possible that the effects of the difference in visual dance between high MS and low MS might have been masked due to the decreased difficulty level as the trial progresses. Additionally, the fact that deviants were only positioned in first, third, fifth and seventh metrical positions (MP) may have made the task easier than in previous research, where deviants were positioned in more diverse positions (MP 1, 4, 5, 7) (Lee *et al.*, 2015). Future research should aim to vary levels of task difficulty so as not to arise adaptation to the task within a trial, but also to not mask a difference between visual stimuli of high MS and low MS.

Overall, this study aimed to examine the probability of MS being a specific factor of dance movement that affects the induction of visual beat which in turn increases the perceptual load on the visual modality and decreases perceptual processing efficiency in auditory modality. Future studies should aim to assess individual differences as well as use tasks with adequate difficulty that would allow investigating the importance of MS in dance for visual beat induction.

References

Arrighi, R., Burr, D., Nazionale, I., & Inoa, A. (2009). *Bordon2012News_Autoimmunity_microRNA-23b keeps TABs on tissue inflammation.pdf*. *9*, 1–7. https://doi.org/10.1167/9.4.25.Introduction

Balasubramanian, S., Melendez-Calderon, A., Roby-Brami, A., & Burdet, E. (2015). On the analysis of movement smoothness. *Journal of NeuroEngineering and Rehabilitation*, *12*(1), 1–11. https://doi.org/10.1186/s12984-015-0090-9

Bidelman, G. M. (2016). Musicians have enhanced audiovisual multisensory binding: experience-dependent effects in the double-flash illusion. *Experimental Brain Research*, *234*(10), 3037–3047. https://doi.org/10.1007/s00221-016-4705-6

Bolger, D., Trost, W., & Schön, D. (2013). Rhythm implicitly affects temporal orienting of attention across modalities. *Acta Psychologica*, *142*(2), 238–244. https://doi.org/10.1016/j.actpsy.2012.11.012

Burger, B., Thompson, M. R., Luck, G., Saarikallio, S. H., & Toiviainen, P. (2014). Hunting for the beat in the body: On period and phase locking in music-induced movement. *Frontiers in Human Neuroscience*, *8*(November), 1–16. https://doi.org/10.3389/fnhum.2014.00903

Chen, Y. C., & Spence, C. (2017). Assessing the role of the "unity assumption" on multisensory integration: A review. *Frontiers in Psychology*, *8*(MAR), 1–22. https://doi.org/10.3389/fpsyg.2017.00445

Grahn, J. A., Henry, M. J., & McAuley, J. D. (2011). FMRI investigation of cross-modal interactions in beat perception: Audition primes vision, but not vice versa. *NeuroImage*, *54*(2), 1231–1243.

https://doi.org/10.1016/j.neuroimage.2010.09.033

Grondin, S., & McAuley, J. D. (2009). Duration discrimination in crossmodal sequences. *Perception*, *38*(10), 1542–1559. https://doi.org/10.1068/p6359

Hove, M. J., Fairhurst, M. T., Kotz, S. A., & Keller, P. E. (2013). Synchronizing with auditory and visual rhythms: An fMRI assessment of modality differences and modality appropriateness. *NeuroImage*, *67*, 313–321. https://doi.org/10.1016/j.neuroimage.2012.11.032

Hove, M. J., Iversen, J. R., Zhang, A., & Repp, B. H. (2013). Synchronization with competing visual and auditory rhythms: Bouncing ball meets metronome. *Psychological Research*, *77*(4), 388–398. https://doi.org/10.1007/s00426-012-0441-0

Lee, K. M., Barrett, K. C., Kim, Y., Lim, Y., & Lee, K. (2015). Dance and music in "gangnam style": How dance observation affects meter perception. *PLoS ONE*, *10*(8), 1–19. https://doi.org/10.1371/journal.pone.0134725

Leman, M. (2007). Embodied Music Cognition and Mediation Technology. *The MIT Press*. https://doi.org/https://doi.org/10.7551/mitpress/7476.001.0001

Lorås, H., Sigmundsson, H., Talcott, J. B., Öhberg, F., & Stensdotter, A. K. (2012). Timing continuous or discontinuous movements across effectors specified by different pacing modalities and intervals. *Experimental Brain Research*, *220*(3–4), 335–347. https://doi.org/10.1007/s00221-012-3142-4

Luck, G., & Sloboda, J. A. (2009). Spatio-temporal cues for visually mediated synchronization. Music Perception, 26(5), 465–473. *Music Perception*, *26*(5), 465–473. https://doi.org/10.1525/mp.2009.26.5.465

Merchant, H., & Honing, H. (2014). Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. *Frontiers in Neuroscience*, *7*(8 JAN), 1–8. https://doi.org/10.3389/fnins.2013.00274

Orlandi, A., Cross, E. S., & Orgs, G. (2020). Timing is everything: Dance aesthetics depend on the complexity of movement kinematics. *Cognition*, *205*(August), 104446. https://doi.org/10.1016/j.cognition.2020.104446

Patel, A. D., Iversen, J. R., Chen, Y., & Repp, B. H. (2005). The influence of metricality and modality on synchronization with a beat. *Experimental Brain Research*, *163*(2), 226–238. https://doi.org/10.1007/s00221-004-2159-8

Repp, B. H., & Penel, A. (2002). Auditory Dominance in Temporal Processing: New Evidence from Synchronization with Simultaneous Visual and Auditory Sequences. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(5), 1085–1099. https://doi.org/10.1037/0096-1523.28.5.1085

Rose, D., Müllensiefen, D., Lovatt, P., & Orgs, G. (2020). The Goldsmiths Dance Sophistication Index (Gold-DSI): A Psychometric Tool to Assess Individual Differences in Dance Experience. *Psychology of Aesthetics, Creativity, and the Arts*, *16*(4), 733–745. https://doi.org/10.1037/aca0000340

Schutz, M., & Lipscomb, S. (2007). Hearing gestures, seeing music: Vision influences perceived tone duration. *Perception*, *36*(6), 888–897. https://doi.org/10.1068/p5635

Sevdalis, V., & Keller, P. E. (2011). Captured by motion: Dance, action understanding, and social cognition. *Brain and Cognition*, *77*(2), 231–236. https://doi.org/10.1016/j.bandc.2011.08.005

Su, Y. H. (2014a). Peak velocity as a cue in audiovisual synchrony perception of rhythmic stimuli. *Cognition*, *131*(3), 330–344. https://doi.org/10.1016/j.cognition.2014.02.004

Su, Y. H. (2014b). Visual enhancement of auditory beat perception across auditory interference levels. *Brain and Cognition*, *90*, 19–31. https://doi.org/10.1016/j.bandc.2014.05.003

Su, Y. H. (2016). Visual tuning and metrical perception of realistic point-light dance movements. *Scientific Reports*, *6*(February), 1–12. https://doi.org/10.1038/srep22774

Thompson, W. F., Russo, F. A., & Livingstone, S. R. (2010). Facial expressions of singers influence perceived pitch relations. In *Psychonomic Bulletin and Review* (Vol. 17, Issue 3, pp. 317–322). https://doi.org/10.3758/PBR.17.3.317

Tsay, C. J. (2014). The vision heuristic: Judging music ensembles by sight alone. In *Organizational Behavior and Human Decision Processes* (Vol. 124, Issue 1, pp. 24–33). https://doi.org/10.1016/j.obhdp.2013.10.003

Wahn, B., & König, P. (2017). Is attentional resource allocation across sensory modalities task-dependent? *Advances in Cognitive Psychology*, *13*(1), 83–96. https://doi.org/10.5709/acp-0209-2