# Visual Cues in Musical Synchronisation

Ian Daniel Colley

A thesis submitted for degree: Doctor of Philosophy

The MARCS Institute for Brain, Behaviour & Development, Western Sydney University

## Acknowledgements

I am grateful to Professor Peter Keller for taking a chance on me, and to Doctor Jennifer MacRitchie and Doctor Manuel Varlet for tolerating my constant state of uncertainty. Somehow, here we are.

# **Statement of Authentication**

The work presented in this thesis is, to the best of my knowledge and belief, original except as acknowledged in the text. I hereby declare that I have not submitted this material, either in full or in part, for a degree at this or any other institution.

Ian Colley

# Declarations

This project was funded by a scholarship from The MARCS Institute for Brain, Behaviour & Development.

<b>TABLE OF CONTENTS</b>
--------------------------

LIST OF FIGURES	<u>iv</u>
LIST OF ABBREVIATIONS	<u>v</u>
ABSTRACT (GENERAL)	<u>vi</u>
INTRODUCTION	1
1. A general introduction to the project	<u>1</u>
2. Musical Timing as Sensorimotor Synchronisation and	_
Two Perspectives on Timing	<u>2</u>
2.1.1. Dynamical Systems Theory and Entrainment	<u>5</u>
2.1.2. Dynamical Systems Theory and Fractality	<u>9</u>
2.2. Information Processing Theory and Internal Timekeepers	<u>13</u>
3. Analytical Differences Between Dynamical Systems and	
Information Processing	<u>16</u>
3.1. Potential connections between dynamical systems and	
information processing	<u><u>19</u></u>
4. Previous studies of visual cues, ancillary movement, and synchroni	sation in
musical contexts	<u>21</u>
5. Overview of Experiments	<u>25</u>
EVERDIMENT 1. The induced of simple and the second section of the	
EXPERIMENT 1: The influence of visual cues on temporal anticipation	27
Abstract	<u></u> <u>29</u>
Abstract and the second s	20
1. Introduction	<u>29</u> <u>36</u>
2. Methods	36
2.1. 1 unicipanis 2.2. Study Design	37
2.2. Shuay Design	37
2.5. Apparatus 2.4 Auditory stimuli	38
2.5 Visual stimuli	
2.6. Virtual conductor	
2.7. Procedure	
3. Data Analysis	
3.1. Drumming SMS	
3.2. Ancillary movements	
3.3. Multiple regression	<u>45</u>
4. Results	<u>45</u>
4.1. Asynchronies and Prediction/Tracking Indices	<u>45</u>
4.2. Motion Capture Analysis	
4.3. Relating movement, prediction, and synchronisation	<u>47</u>
5. Discussion	<u>48</u>
6. Conclusion	<u><u>56</u></u>
Acknowledgements	<u>57</u>

EXPERIMENT 2: The influence of a conductor and co-performer	
on auditory motor synchronisation, temporal prediction and	
ancillary entrainment in a musical drumming task	<u>58</u>
Abstract	<u>59</u>
1. Introduction	<u>60</u>
2. Methods	<u>64</u>
2.1. Participants	<u>64</u>
2.2. Apparatus	<u>65</u>
2.3. Auditory stimuli	<u>65</u>
2.4. Visual stimulus	
2.5. Procedure	<u>67</u>
2.6. Design	<u>68</u>
3. Data analysis	
3.1. Drumming data	<u>69</u>
3.2. Ancillary movements	$\frac{70}{72}$
4. Results	
4.1. Synchronisation drumming: Asynchrony	
4.2. Synchronisation arumming: Anticipatory timing	
4.5. Asynchrony as a function of anticipatory timing	
4.4. Synchronisation of ancittary movement	<u>70</u>
4.5. Retaining unclitary and instrumental movements	<u>70</u> <u>80</u>
<b>5. Discussion</b>	<u>00</u>
EXPERIMENT 3: Does movement amplitude of a co-nerformer	
affect individual performance in musical synchronisation?	89
Abstract	<u></u> <u>90</u>
1. Introduction	<u>91</u>
2. Methods	
2.1. Participants	
2.2. Design	
2.3. Auditory stimuli	<u>95</u>
2.4. Visual stimuli	<u>95</u>
2.5. Apparatus	<u>96</u>
2.6. Procedure	<u>97</u>
3. Data Analysis	<u>98</u>
3.1. Drumming analysis	<u>98</u>
3.2. Motion capture analysis	<u>99</u>
3.3. Equivalence test	<u>100</u>
4. Results	<u>102</u>
4.1. Asynchrony	<u>102</u>
4.2. <i>P/T Index</i>	<u>103</u>
4.3. DFA	<u>103</u>
4.4. Coherence	<u>104</u>
5. Discussion	<u>106</u>
CENEDAL DISCUSSION	110
1 Deview of the Desults	<u>112</u> 112
<ol> <li>A Content of the Acoustic Acoustic</li></ol>	<u>114</u> 114
2. Elimitations and Future Directions	<u>114</u>
CONCLUSION	<u>117</u>

	 <b>REFERENCES</b>
APPENDIX	 APPENDIX
1. Appendix A	 1. Appendix A

## LIST OF FIGURES

## Introduction

- 1. Figure 1.1. Signals exhibiting statistical noise
- 2. Figure 1.2. A depiction of asynchrony and correction in a tapping task
- 3. Figure 1.3. A depiction of phase difference in a pendulum task
- 4. Figure 1.4. A depiction of how timing mechanisms...

## **Experiment 1**

- 5. Figure 2.1. Diagrams of the visual cue conditions
- 6. Figure 2.2. Mean absolute asynchronies for each expertise group...
- 7. Figure 2.3. Mean P/T Index for each expertise group...
- 8. Figure 2.4. Average standard deviation of movement
- 9. Figure 2.5. Scatterplots showing the relationship between...

## **Experiment 2**

- 10. Figure 3.1. A schematic of the three experimental conditions
- 11. Figure 3.2. The mean absolute value of asynchronies...
- 12. Figure 3.3. The mean P/T index across the three visual cue conditions
- 13. Figure 3.4. The maximum coherence values...
- 14. Figure 3.5. Mean coherence values from each measured frequency bin...
- 15. Figure 3.6. The 95% confidence intervals of difference scores...
- 16. Figure 3.7. Scatterplots showing the relation between asynchrony...and coherence

## **Experiment 3**

- 17. Figure 4.1. A schematic of the three moving visual cues
- 18. Figure 4.2. The mean asynchrony scores expressed in milliseconds
- 29. Figure 4.3. The mean P/T Index scores...
- 20. Figure 4.4. The mean cross-spectral coherence scores...
- 21. Figure 4.5. The mean DFA scores...
- 22. Figure 4.6. The equivalence bounds and corresponding TOST results...

## Appendix

23. Appendix A. A sample of the music used as auditory stimuli

# LIST OF ABBREVIATIONS

SMS = sensorimotor synchronisation CC = cross-correlation IOI = inter-onset interval ITI = inter-tap interval

#### ABSTRACT (GENERAL)

Although music performance is generally thought of as an auditory activity in the Western tradition, the presence of continuous visual information in live music contributes to the cohesiveness of music ensembles, which presents an interesting psychological phenomenon in which audio and visual cues are presumably integrated. In order to investigate how auditory and visual sensory information are combined in the basic process of synchronising movements with music, this thesis focuses on both musicians and nonmusicians as they respond to two sources of visual information common to ensembles: the conductor, and the ancillary movements (movements that do not directly create sound; e.g. body sway or head nods) of co-performers. These visual cues were hypothesized to improve the timing of intentional synchronous action (matching a musical pulse), as well as increasing the synchrony of emergent ancillary movements between participant and stimulus. The visual cues were tested in controlled renderings of ensemble music arrangements, and were derived from real, biological motion. All three experiments employed the same basic synchronisation task: participants drummed along to the pulse of tempo-changing music while observing various visual cues. For each experiment, participants' drum timing and upper-body movements were recorded as they completed the synchronisation task. The analyses used to quantify drum timing and ancillary movements came from theoretical approaches to movement timing and entrainment: information processing and dynamical systems.

Experiment 1 focused on the influence of the conductor by comparing three types of visual stimuli: a prototypical conductor gesture (derived from motion capture of real conductors), a simple visual metronome devoid of biological motion, and a still image. There was also a comparison of two participant groups: musicians and nonmusicians. The conductor was expected to be the superior visual cue due to the high rates of acceleration and

vi

velocity along its trajectory. The musicians were expected to perform better overall but show the same pattern of results as nonmusicians since visuo- and audio-motor entrainment appears to be a general, not specialised process. Results indicated that the conductor cue indeed yielded lower asynchronies and higher temporal prediction. This was true for both groups, though musicians performed better overall. For the motion capture recording, the nonmusician group showed increased head movement fluctuations with the conductor, while the musicians showed very little movement overall.

Having established that nonmusicians can complete the musical task devised for this thesis, and that they are sensitive to conductor gestures, Experiment 2 tested the combined influence of a conductor and co-performer using a general sample in a dyadic synchronisation task. The manipulation this time was the arrangement of dyads, so that they could see 1) each other and the conductor; 2) the conductor but not each other; or 3) one person could only see the conductor, while the other person could only see their partner. Asynchrony was highest in the third condition, suggesting observing the same temporal information is preferable to an asymmetrical arrangement. The experiment also showed the presence of visuo-motor entrainment among partners' ancillary movements, and that higher ancillary synchrony was associated with lower drumming asynchrony at certain frequencies of movement.

Given the evidence that a live co-performer can influence synchronisation, Experiment 3 removed the conductor to test the effects of a virtual co-performer only. Similar to the virtual conductor, the virtual co-performer was made using motion capture recordings of previous high-performing participants. The manipulation was the amplitude of the coperformer's movements (natural range of motion, doubled range, tripled range, or a still image control), as previous research has shown mixed results regarding the effect of visual stimulus amplitude on synchronisation. Therefore, testing movement amplitude was intended to address an unresolved question. However, the only significant effect was between the nonmoving control image and all moving stimuli on the participant's ancillary movement synchrony, showing again the presence of visuo-motor entrainment. There was no effect on drumming asynchrony.

By using methods from traditionally opposed theoretical viewpoints (information processing and dynamical systems), these experiments demonstrate that both perspectives can contribute to our understanding of musical timing. More specifically, these findings provide evidence for the phenomenon of visuo-motor entrainment in musical synchronisation, suggesting it may be an important aspect of ensemble music performance. The results also show empirically that a conductor can reduce asynchronies between co-performers even in the context of laboratory drumming task, and that this reduced asynchrony may be explained by an increase in temporal prediction when observing the conductor. A live co-performer also seems to influence musical timing, but a virtual one does not, at least in the general population. Overall, this thesis shows that basic musical timing is a common ability that is facilitated by visual cues in certain contexts, and that emergent ancillary movements and intentional synchronous movements in combination may best explain musical timing and synchronisation.

#### **INTRODUCTION**

## 1. A general introduction to the project

Ensemble music performance is a complex, widespread, and uniquely human behaviour that represents the union of many cognitive, motor, and social skills. Empirical investigations of music cognition and action—sometimes called empirical musicology (Clarke & Cook, 2004)—have started to identify the specific physical, biological, and psychological processes underlying this high-level human activity. Such research has had applications in music pedagogy (Luck & Sloboda, 2007), as well as more general applications to mood regulation and mental health (Garrido, Eerola, & K., 2017; Garrido & Schubert, 2015a, 2015b; Saarikallio, 2008), and physical therapy for neurodegenerative disorders (Benoit et al., 2014; M.J. Hove & Keller, 2015; M.J. Hove, Suzuki, Uchitomi, Orimo, & Miyake, 2012; Nombela, Hughes, Owen, & Grahn, 2013). Also, because music encompasses both motor skills and cognitive functions, and because participating in musiceither through performance as a trained musician or through listening as an audience member—is a ubiquitous aspect of human culture, such research lends itself to unravelling the foundational processes of human behaviour in a controlled, but ecologically valid setting. In short, music can be both a tool for improving skills, and a method of understanding behaviour<sup>1</sup>, depending on research aims and design.

The body of work presented here runs primarily along basic lines of research and examines the phenomenon of human movement coordination in the context of musical synchronisation using both nonmusicians and trained musicians. More specifically, this series of experiments looks at movement timing and sensorimotor synchronisation as they are influenced by dynamic visual cues that might be seen in ensemble music. Although musical

<sup>&</sup>lt;sup>1</sup> Music can of course also be an art form or creative pursuit that stands on its own. The research discussed here however is focused on interdisciplinary empirical musicology and experimental psychology.

activities are often enjoyed independently by individuals, music is very often social (D'Ausilio, Novembre, Fadiga, & Keller, 2015), as seen in the enduring popularity of ensemble practice and performance for most of human history (Hodges, 1996).

From a psychological standpoint, joint music production is especially interesting given the inter-dependence of co-performers in their attempt to maintain temporal structure and synchrony (Keller & Burnham, 2005; R. A. Rasch, 1979; Wing, Endo, Yates, & Bradbury, 2014) as well as the multi-sensory nature of this co-performer relationship; while processing their own proprioceptive and tactile states, ensemble musicians are also involved in a continuous exchange of auditory and visual information. Their actions are informed by the sounds they produce, and by the movements they see in the ensemble. These factors combine to create a dynamic network of musical communication that showcases the interconnectivity of human movement and temporal processing.

The study of temporal processing dates back to over a century (Stevens, 1886), but the basic method of having subjects tap their index finger and recording inter-response intervals has remained largely the same (Michon, 1967; Repp, 2005; Repp & Su, 2013). Research on the dynamics of human movement, however, has thrived more recently with the advent of sophisticated motion tracking technology (Burger, Thompson, Luck, Saarikallio, & Toiviainen, 2014; Richardson, Marsh, Isenhower, Goodman, & Schmidt, 2007; Shockley, Baker, Richardson, & Fowler, 2007). These two fields and their associated methods rarely cross paths due to theoretical disagreements (Delignières & Torre, 2011). This matter is discussed later, but one goal of this thesis was to analyse music-related movement and timing using methods from both schools of thought.

## 2. Musical Timing as Sensorimotor Synchronisation and Two Perspectives on Timing

The coordination of movements with sensory inputs is known as sensorimotor synchronisation (SMS), and describes a wide range of periodic human motor behaviour such

walking (H. Y. Chen, Wing, & Pratt, 2006; Nessler & Gilliland, 2009) and conversing (Shockley et al., 2007). SMS can also be described as the coordination of internal timing (the so-called *time sense* (Stevens, 1886)) and external timing (e.g. events perceived in the environment). It is unsurprising then that SMS has been used as a theoretical basis for understanding music performance, as most musical styles are rigorously structured by rhythms, mediated by multi-sensory information (such as a metronome, auditory feedback, or a conductor's baton), and have a goal of achieving synchronised sounds.

Most people exhibit simple SMS (e.g. when tapping along to a steady beat at a concert) automatically (Repp, 2001) and with apparent ease. However, the ability to match movement timing to periodic stimuli is not perfect due to biological noise (Cohen & Sternad, 2009; Pressing & Jolley-Rogers, 1997), and therefore not always successful (M.J. Hove, Spivey, & Krumhansl, 2010). Furthermore, synchronising with a beat appears to be absent in a small population of "beat deaf" individuals (Phillips-Silver et al., 2011). Such timing deficiencies suggest that SMS is not always an automatic process. Furthermore, individual differences in rhythmic abilities and SMS (Grahn & Schuit, 2012; Iversen & Patel, 2008) also suggest that periodic sensory inputs do not necessarily result in synchronous motor outputs; in other words, SMS is clearly more complicated than a 1:1 relationship between bodily and environmental rhythms. Thus, SMS has rightfully become a widely studied phenomenon (Repp, 2005; Repp & Su, 2013), and is a valid starting point for investigating musical synchronisation. Questions of how SMS is realized and the extent to which individuals can control the timing of their movements, however, have diverged into two theoretical approaches. To illustrate, consider the following examples.

First, the ability to control one's movements accurately and precisely in musical contexts is not unique to musicians. Indeed, many studies of synchronisation and timing recruit non-musicians and find that timing capabilities are generally stable (Repp, 2008)

although musicians do tend to perform with greater accuracy and precision (Repp, 2010). Second, the phenomenon of motor synchronisation with musical stimuli is easily observed anecdotally as well, as in audience members nodding their heads or tapping their toes in time to a performance (Witek et al., 2017).

While based on similar observable behaviours (rhythmic movements), these two examples are notably different in the extent to which the activities described are intentional and goal-directed. The former—participants in synchronisation studies—is a situation in which individuals are tasked with timing their actions to match a stimulus. Thus, the participants are aware of their timing and have an objective. The latter—moving in synchrony with an observed performance—describes a spontaneous behaviour. The audience member's attention is given to the performer, not to his or her own rhythmic movements, and there is probably no objective to synchronise. These examples, therefore, are representative of two different, often juxtaposed theories of how periodic motor timing is realized: information processing theory and dynamical systems theory.

The information processing perspective posits that movement timing is driven by the *intention* to move, and that units of time are represented in memory. In order to execute periodic movements such as finger tapping, there must be a central representation of the time period separating each motor event (Stevens, 1886; Vorberg & Wing, 1996; Wing & Kristofferson, 1973). Noise in the central representation and in the motor system adds variability to the timing process (Wing & Kristofferson, 1973), but feedback and error correction mechanisms are typically able to maintain consistency in the timed behaviour (H. Y. Chen et al., 2006; Praamstra, Turgeon, Hesse, Wing, & Perryer, 2003; Repp, 2002b; Repp & Moseley, 2012). In the case of synchronisation with a target stimulus (as opposed to self-paced tapping), information processing proponents would argue that periodic movements are organized according to an apparent recurring target such as a musical pulse, and errors

associated with one target are accounted for at the next target (van der Steen & Keller, 2013). Given the importance of these targets or "events," timing in the information processing school is often referred to as *event-based* (Delignières & Torre, 2011; Repp & Steinman, 2010; Torre & Delignières, 2008). In this line of thought, limitations in timing could be due to factors such as limited speed of processing at very high event rates (Repp, 2003b), and limited attentional resources in complex environments such as ensemble performances (Inhoff & Bisiacchi, 1990).

Dynamical Systems explanations of movement timing come from a much larger subfield of physics that seeks to explain how systems of any sort—population growth, chemical reactions (Strogatz, 1995), and audience applause (Néda & Ravasz, 2000) for example change and self-organize over given time periods. Because dynamical systems are selforganizing, any semblance of order in this system does not come from an executive or controlling force, but from physical restrictions in the system. In the case of repetitive timing activities, periodicity is not the goal but the product of movement because the cells, bones, and muscles involved must organize in an oscillatory manner. For this reason, timing in the dynamical systems school is often described as *emergent* (Delignières & Torre, 2011; Repp & Steinman, 2010; Zelaznik, Spencer, & Ivry, 2002) because it is simply a by-product of more fundamental processes that are, according to strong proponents of dynamical systems, not intentional (Treffner & Turvey, 1993). Limitations in timing accuracy and consistency may be due to physical limitations such as weak neuromuscular coupling, or disrupted sensorimotor coupling (Miyata, Varlet, Miura, Kudo, & Keller, 2017; Richardson et al., 2007).

## 2.1.1 Dynamical Systems Theory and Entrainment

A common theoretical foundation for synchronisation in the dynamical systems school is *entrainment*. Entrainment is the spontaneous synchronisation of periodic behaviours between two or more systems and was first studied in clocks in the 17<sup>th</sup> century by Christiaan Huygens (described in Strogatz & Stewart, 1993). It was observed that pendulum clocks placed next to each other on a shelf would begin ticking in synchrony, despite initially being set out of sync. This led to the proposition that entrainment requires a *coupler*; that is, in order for two periodic behaviours to become synchronised, the systems producing the behaviour must be *coupled* by contact through a common medium, which is the shelf in the clock example.

Since this discovery, entrainment has become a widely studied phenomenon, observable in many types of periodic behaviour, across many fields of science, and over a wide range of timescales. Circadian rhythms entrain to the time of day (Komin, Murza, Hernandez-Garcia, & Toral, 2011); fireflies blink in synchrony by entraining to each other (Strogatz & Stewart, 1993); and—most relevant to ensemble synchronisation—periodic human movements can entrain to periodic exogenous stimuli, such as the pulse in music or the movements of others (Large, 2008; Oullier & Kelso, 2009). In the case of ensemble synchronisation, two co-performers can be thought of as two clocks. Their "ticking" is the production of tones based on an established pulse. Assuming the co-performers are coupled by hearing and/or seeing each other, their tone onsets will presumably synchronise, as each process exerts influence on the other. If sensory information is cut off, the co-performers will still act periodically independently, but likely fall out of sync with each other due to a lack of coupling.

A central concept in SMS as an entrained dynamical system is that synchronous behaviour is spontaneous and sometimes even unintentional. Spontaneous entrainment has been observed in tasks in which participants are instructed to freely swing a handheld pendulum. The presence of a periodic visual stimulus influences the freely swinging pendulum, as indicated by phase locking of the pendulum to the stimulus (Demos, Chaffin,

Begosh, Daniels, & Marsh, 2012; Schmidt, Richardson, Arsenault, & Galantucci, 2007; Varlet, Bucci, Richardson, & Schmidt, 2015). Unintentional entrainment refers to phase locking with a stimulus, despite instructions to maintain some other oscillation frequency (Varlet, Coey, Schmidt, & Richardson, 2012). These phenomena are indicative of the readiness with which humans entrain movements to exogenous periodic stimuli. The more general examples listed above (e.g. fireflies blinking together) are indicative of the ubiquity and naturalness of entrainment, further suggesting that it is a valid theoretical consideration for explaining musical behaviours.

Importantly, the strength of the coupling, and thereby the influence of an environmental rhythm on an individual's movements, is dependent on a number of factors. For example, coupling strength will likely increase with the amount of visual information about the oscillator with which the individual is synchronizing. For instance, visually tracking a moving stimulus compared to fixating on a stable point just above the midpoint of the stimulus trajectory was found to increase the strength of entrainment in a pendulum task (Varlet, Bucci, et al., 2015), possibly because tracking the stimulus increases pickup of information about the stimulus trajectory. Similarly, the salience of a stimulus as measured by its amplitude can affect coupling (Varlet, Coey, et al., 2012). Again, this could be due to the availability of more information about the movement. Moreover, removing all visual access between two partners in a dyadic scenario causes synchronous behaviour to dissipate, while subsequently restoring visual access also restores the partners' synchronisation (Oullier & Kelso, 2009).

Providing multimodal information can also strengthen coupling (Elliott, Wing, & Welchman, 2010). For instance, individuals synchronised their pendulums with a stimulus better when the stimulus contained both auditory and visual information (Armstrong & Issartel, 2014). This study also found an advantage for continuous over discrete visual

information, and discrete over continuous auditory information in synchronisation, a finding that has been reported consistently in the synchronisation literature (Bishop & Goebl, 2014; Chen, Repp, & Patel, 2002; Grahn, 2012; Repp & Su, 2013).

Overall, studies of human-human and human-to-stimulus entrainment have shown that individuals coordinate their motor timing with periodic environmental timing both spontaneously and unintentionally. Visually tracking a stimulus strengthens coupling, but entrainment effects are still found when the stimulus is observed peripherally (Richardson et al., 2007). Lastly, there seems to be an advantage when coupling with continuous visual stimuli, but this effect is enhanced when a discrete auditory stimulus is added to emphasize the turnaround point in a movement cycle (Varlet, Marin, Issartel, Schmidt, & Bardy, 2012; Zelaznik & Rosenbaum, 2010).

When attempting to synchronise in a music ensemble, a performer's attention is likely to shift from focusing directly on a co-performer whose timing information is crucial at a given point in the music, to focusing on the score or conductor, at which point peripheral views of co-performers could help maintain a steady phase among players via entrainment of ancillary movements (movements such as upper-body sway that are not causally linked to producing sound). Measuring the extent to which ancillary movements and conductor kinematics play a role in shaping dynamics of participant movements, and testing whether individuals can obtain temporal information from the ancillary movements of others is one goal of the experiments described below.

Ancillary movements, as opposed to instrumental movements (movements that cause the production of sound, e.g. violin bow strokes) are of interest because these movements represent a more generalizable type of motion. Unless research participation is limited to a homogeneous group of musicians, different types of musical experience could confound any effects of instrumental movements, given that instrumental movements differ per family of

instrument. Furthermore, the instrumental movements of a given musician in an ensemble may not be peripherally visible to all other musicians in the ensemble. Some degree of grossmotor upper-body sway is common and likely visible throughout an ensemble. Thus, although musicians certainly rely on the instrumental movements of co-performers in the immediate vicinity, this project focuses on ancillary movements, as they contribute to a more widely accessible visual environmental rhythm.

However, it is important to note that ancillary movements do not necessarily represent a one-to-one relationship between somatic cues and a psychological sense of rhythm. Musicians may show idiosyncratic patterns of ancillary movements in addition to patterns of movement that are common across individuals—and therefore might represent commonalities in the music being performed—as has been demonstrated in clarinettists (Wanderley, 2002; Wanderley, Vines, Middleton, McKay, & Hatch, 2005). Furthermore, it should be noted that *ancillary movements* as defined here (i.e. not essential to the production of sound) have also been referred to as *nonobvious performer gestures*, implying that this category of movements actually can affect sound (Wanderley, 1999). For the sake of this project, a distinction is drawn between movements that directly control musical sounds (namely the arm/hand) and those that are distally related to musical sounds (namely head movements).

## 2.1.2. Dynamical Systems Theory and Fractality

Another approach to studying the dynamics of complex systems is fractal analysis (Delignières, Torre, & Lemoine, 2008; Madison, 2004; Riley, Bonnette, Kuznetsov, Wallot, & Gao, 2012; Stadnitski, 2012). A system is fractal if it is self-similar at multiple scales, meaning a part will bear some resemblance to the whole from which it came. In terms of human movement timing, variability in performance, if it is fractal, will resemble itself at different timescales (e.g. variance over a span of seconds will resemble variance over a span of minutes). Whereas entrainment describes how two systems come to be coupled and how their trajectories are related (Strogatz & Stewart, 1993), fractal analysis can identify supposed<sup>2</sup> long-range correlations within a system (Duarte & Zatsiorsky, 2001; Torre, Balasubramaniam, Rheaume, Lemoine, & Zelaznik, 2011). Long-range correlations are believed to indicate the extent to which the system is self-organizing (Bak, Tang, & Wiesenfield, 1987; Van Orden, Holden, & Turvey, 2003), which is considered a reflection of the system's efficiency; no long-range correlation indicates a random series of events, very strong or persistent long-range correlations indicate rigidity and order, and a middle ground likely reflects optimal efficiency governed by self-organization (Stadnitski, 2012).

One method by which long-range correlations are identified is by measuring *statistical noise*. In this sense, *noise* refers to patterns of variability in a time series rather than complete randomness or measurement error. Analysis of statistical noise has been applied to numerous complex systems, including human movement timing. By plotting some measure of movement as a time series (such as centre of pressure while standing, time between footsteps while walking, or pendulum turnaround times), one can analyse the nature of the variability, or fluctuations over time and identify the presence of white noise, pink noise, or Brownian noise (see **Figure 1.1**). White noise indicates no long-range correlations (randomness) and is the product of a statistically stationary time series. Any influence between successive events will not persist in subsequent observations, and therefore have no influence on the future state of the system. Pink noise, also called1/*f* noise, indicates an optimal level of long-range correlations<sup>2</sup> in a system. This means that events will be similar at very distant points in the time series, and so a given event will likely have significant influence on future events for well past a few observations due to the nature of self-organizing dynamical systems (Stadnitski, 2012; Wagenmakers, Farrell, & Ratcliff, 2004;

<sup>&</sup>lt;sup>2</sup> Note that the concept of long-range correlations has been contested. While the basic shape and dynamics of a time series might genuinely adhere to pink-noise structure, the serial correlations in the series may in fact be short-range in nature (Madison & Delignéres, 2008; Pressing, 1999b)

Ward & Greenwood, 2007). *Brownian noise*, sometimes called *red noise*, indicates that longrange correlations are *super-persistent*, meaning any change in the system will influence future events necessarily, perhaps as a result of intention and cognitive control (Lacour, Bernard-Demanze, & Dumitrescu, 2008).

In terms of human movement coordination, pink noise could represent an ideal timing mechanism with which an individual is able to adapt his or her movements according to both inter- and intra-personal demands that might arise, and thereby settle into a stable state more readily. Indeed, pink noise correlations have been identified in healthy gait timing (Lamoth, Lummel, & Beek, 2009), standing balance (Blázquez, Anguiano, Saavedra, Lallena, & Carpena, 2009; Duarte & Zatsiorsky, 2001), and rhythmic tapping (Delignières et al., 2008; Hennig et al., 2011). Apart from the ubiquity of pink noise in stable systems, the significance of pink noise in complex system organization is further suggested by its absence in abnormally functioning or unstable systems, such as balance in older adults (Wang & Yang, 2012), pathological heartbeats (Peng, Havlin, Stanley, & Goldberger, 1995), and gait timing in Parkinson's Disease (Hove et al., 2012). In terms of music, listeners tend to prefer "humanized" music with slight timing errors that exhibit pink noise over perfect timing or random timing errors (Hennig et al., 2011). Regarding musical movement, synchronisation over very long music sequences (> 1000 pulses) showed 1/f correlations (Hennig et al., 2011), as did expressive body sway in trombonists (Demos, Chaffin, & Kant, 2014), and pulse times in piano performances (Rankin, Large, & Fink, 2009).



*Figure 1.1.* Signals exhibiting statistical noise. From top to bottom: white, pink (1/f), and Brownian noise (or red noise). Note that the Brownian noise signal strays further from a stationary mean than pink or white noise, reflecting the ongoing change in Brownian noise systems. This is also seen in the wider range of values on the y axis, showing that a Brownian noise system will often move beyond the scale of a white or pink noise system. Scales are arbitrary values.

Taken together, these findings suggest human movement timing naturally exhibits stable dynamical system properties, as seen in temporal organization that adheres to 1/f correlational structure. Body movements adhering to this structure in musical contexts likely allow a person to shift between automatic temporal processing during periods when a synchronisation stimulus is predictable and regular, and cognitively mediated processing when the stimulus is unsteady (e.g. a tempo change) and therefore cognitively demanding

(Colley, Keller, & Halpern, 2017). In other words, pink-noise structures in the human motor system may be utilized in musical timing to adjust coordination according to both expressive (Demos et al., 2014; Rankin et al., 2009) and rhythmic (Hennig et al., 2011; Ruiz, Hong, Hennig, Altenmuller, & Kuhn, 2014) demands. The studies presented later have measured the noise colour of participants during musical synchronisation in order to describe how pink noise structure in ancillary movements is influenced by visual cues in music. Given the prior research on entrainment to visual cues, one would expect visual cues in music contexts to affect noise colour. Specific hypotheses are discussed with each experiment.

## 2.2. Information Processing Theory and Internal Timekeepers

The information processing theory of motor timing begins with the presupposition that segments of time are represented by central timekeepers, a mental construct of a time interval. Such an assumption is readily applied to musical time-keeping, as people often represent musical beat intervals by deliberately counting sub-vocally. Central timekeepers are most commonly tested using finger-tapping tasks in which participants are instructed to synchronise taps with a periodic stimulus such as a metronome or attempt to tap at regular intervals at a self-determined pace. Wing and Kristofferson (1973) developed a model of this process in self-paced tapping (i.e. not synchronisation with external pacing events) using inter-tap intervals (ITIs) to estimate the activity of the proposed internal clock. The Wing-Kristofferson model has since been expanded to account for synchronisation timing by including error correction parameters based on asynchronies between tap intervals (ITIs) and inter-onset intervals (IOIs, see Figure 2) of metronome clicks (Vorberg & Schulze, 2002; Vorberg & Wing, 1996). These information processing models suggest that a given ITI (and by extension the corresponding internal representation of time) is serially dependent on the preceding tap, such that inaccurately long intervals are followed by shorter intervals to make up for the error, and vice versa. Information processing models such as the WingKristofferson model have also established the practice of using autoregressive modelling to predict temporal intervals from preceding (Jacoby & Repp, 2012; Pressing & Jolley-Rogers, 1997; Vorberg & Hambuch, 1978; Wing & Beek, 2002).

When engaged in a timing task, individuals usually exhibit *adaptive timing*. This refers to the adjustment of taps (or some other repetitive behaviour) to account for errors relative to a target time such as a musical pulse. Adaptive timing for isochronous sequences at a local level—known as *phase correction*— is believed to be largely automatic (Keller, 2013; Konvalinka, Vuust, Roepstorff, & Frith, 2010; Repp & Keller, 2004; van der Steen & Keller, 2013). Adaptive timing at larger timescales, such as adjusting taps to match tempo-changing sequences, is known as *period correction*. This requires more attentional resources, but is mediated by internal representations of time similar to phase correction (Repp & Keller, 2004). Parameter estimates of the extent to which individuals exhibit adaptive timing (Repp, Keller, & Jacoby, 2012; van der Steen & Keller, 2013) are based largely on studies of phase and period correction responses, or, respectively, how quickly a person responds to an experimentally added phase shift (i.e. a disruption of sequence regularity) in a target time series and how persistent that correction is as the perturbation continues for subsequent intervals (Repp, 2002b, 2005).

However, for tap times or other timed actions to be processed and adjusted online fast enough to maintain synchrony, there is likely another facilitating cognitive process in addition to the error correction described above. Indeed, several studies of SMS have pointed towards the importance of *anticipatory timing* during synchronisation. Anticipatory or predictive timing refers to the prediction of upcoming time intervals, and is likely based on auditory imagery of what will happen next in a sequence (Colley et al., 2017; Keller, 2014; Keller & Appel, 2010; Pecenka & Keller, 2009b; Repp & Su, 2013). This is especially important in the case of ensemble performance, in which the IOIs of musical beats are often intentionally irregular in order to add expressive nuances to the music (Rankin et al., 2009; Repp, 1998; Shaffer, 1984), or to change the tempo of the music, or simply due to human error. Michon first proposed a quantitative measure of "time tracking" (Michon, 1967), and put forth the idea that timing in musical behaviours is based on previous temporal events. This idea underlies measurements of anticipatory timing in musical contexts. Originally, this was quantified by cross-correlating the time series of motor events/taps with the time series of target auditory events, and comparing that coefficient to the cross-correlation coefficient between motor events, and the lag-1 target event series (Colley et al., 2017; Rankin et al., 2009; Repp, 1998). If participants are tracking target intervals, their tap sequence will correlate with the lag-1 target series, whereas if they are predicting then their tap sequence should resemble the original target series. Differencing these two coefficients produces a prediction-tracking (P-T) index, which is positive to the extent that an individual is predicting (Pecenka & Keller, 2011; Repp, 2002a). The experiments that form this thesis first used this method of measuring predictive timing. However, a similar but more sophisticated method of assessing predictive timing was proposed and used for the latter two experiments. The details of this method are described in the experiments.

Such measurements have shown that most people do tend to predict rather than track IOIs, and that the extent of prediction is positively correlated with SMS performance (Colley et al., 2017; Pecenka & Keller, 2009a, 2009b) in tempo changing sequences. However, SMS in "trackers" can be improved somewhat in interpersonal tapping tasks when a tracker is paired with a good predictor (Pecenka & Keller, 2011), suggesting that interactions with co-performers can be beneficial to synchronisation under certain conditions. Evidence from tapping tasks also suggests that anticipatory timing can be cued, as seen in a study of anticipatory phase correction (Repp & Moseley, 2012). Specifically, when alerted to an

upcoming phase perturbation, participants were quicker to adjust to the shifted sequence after it occurred.

Although the existence of predictive mechanisms in motor timing is well documented, even at the neural level (Penhune & Zatorre, 1998), little work has been done on what external factors can facilitate anticipatory timing in periodic actions. An established, isochronous pulse is easily predicted, and adaptive timing optimizes synchronisation with such sequences. However, pulse sequences in music are rarely isochronous, but instead shaped by tempo fluctuations according to a composer's instructions and the interpretations of performers or a conductor. In tempo-changing passages, anticipatory timing would be particularly important, and the conductor should alert performers to upcoming changes in the interest of reducing asynchronies. Indeed conductors are often trained to emphasise a beat or tempo change ahead of the ensemble, presumably in an attempt to indicate deviations from a regular beat. Part of this project will assess anticipatory timing can be improved by the presence of an informative virtual conductor or co-performer.

## 3. Analytical Differences Between Dynamical Systems and Information Processing

The information processing approach to understanding timing typically measures asynchrony and adaptive parameters by comparing cumulative times of each tap to the corresponding target event, resulting in a series of discrete values that correspond to each action or event (see **Figure 2**). The primary dependent measure of synchrony in experiments motivated by dynamical systems theory is *relative phase*, which is conceptualised as follows: the range of participant's and stimulus' movements are converted to degrees in a circle. Then, one can measure the angular phase difference between the movement cycles at any given time point within a cycle, with the starting point of the action at 0°, the halfway point at 180°, and the completion at 360° (see **Figure 3**).



Figure 2. A depiction of asynchrony and correction in a tapping task. The IOI sequence shown here is slowing down (the time intervals are enlarging). The tap sequence is initially leading the IOI sequence, but shows one very large ITI (600 ms) to account for previous erroneously short intervals error. The asynchrony is calculated after each pulse interval as the difference between the cumulative time of ITIs and IOIs. The mean asynchrony would be the average of the asynchrony series (bottom row; 24 ms mean asynchrony). Although taps can lead or lag a pulse, creating negative and positive asynchronies, asynchrony is reported in absolute values throughout this project.



Figure 3. A depiction of phase difference in a pendulum task. The starting position (the left side of the pendulum trajectory) corresponds to  $0^{\circ}$ , the halfway point to  $90^{\circ}$  (or  $270^{\circ}$  on the return swing), and the right extreme to  $180^{\circ}$ . Note that the pendulum does not need to move in a full circle to be measured angularly. Rather, the completion of a movement cycle from starting point, to the opposite extreme, back to the starting point creates a cycle of  $360^{\circ}$ . The angular difference between two pendulums at some point during this trajectory is the phase difference, marked by the dotted curves. Here, P2 is lagging behind P1 consistently by about  $20^{\circ}$ .

One complication when interpreting these measurements in a general theory of timing is that information processing focuses on between cycle processes (planning for the next event or responding to a previous one) whereas dynamical systems measures within-cycle dynamics (what happens while the timed action is unfolding). It is possible, however, that some actions contain both emergent and event-based timing properties (Repp & Steinman, 2010). Other studies of event-based and emergent timing failed to support simultaneous timing mechanisms, but did find evidence that individuals may switch between event-based (cognitive) and emergent (dynamical) timing depending on task constraints (Delignières & Torre, 2011; Huys, Studenka, Rheaume, Zelaznik, & Jirsa, 2008; Varlet, Marin, et al., 2012). Another attempt at joining the two theories examined the within- rather than between-cycle dynamics of tapping; tapping is generally thought of as a discrete action and therefore is normally subject to between-cycle analysis. Results showed that people tend to move their fingers slowly away from a target surface, then dwell at the peak before moving rapidly towards the surface just before the time of the target pulse (Balasubramaniam, Wing, & Daffertshofer, 2004). This asymmetry in movement for discrete timing seems to facilitate faster error correction by providing more perceptual information about the arrival time of an effector (Elliott, Welchman, & Wing, 2009) and allotting processing time during the dwell position (Balasubramaniam et al., 2004). Importantly for a conductor in a music ensemble, this sort of velocity profile might also convey timing information to others.

## 3.1 Potential connections between dynamical systems and information processing

Although dynamical systems and information processing theories of motor timing were once considered mutually exclusive alternatives, more recent approaches have recognized that both considerations are needed to explain human movement and timing. For example, dynamical systems approaches provide more thorough accounts of the physical constraints affecting the motor system (Wing & Beek, 2002). Furthermore, the dynamical systems principle of self-organization overcomes the homunculus problem associated with cognitive control theories (i.e. who controls the controller? [Logan, 2003]). On the other hand, information processing approaches acknowledge, and have begun to explain the effects of cognition, such as learning, memory load, and other non-motor, centrally mediated individual differences (Jacoby, Jakobi, Lieder, Tishby, & Ahissar, 2013; Maes, Wanderley, & Palmer, 2015; Pecenka & Keller, 2009a; Ragert, Schroeder, & Keller, 2013). Some researchers have claimed that this apparent dichotomy is actually a continuum from selforganization to cognitive control at the extremes (Delignières & Torre, 2011). Either way, both the dynamical systems and information processing schools—while based on some substantiated and oppositional differences—are important for constructing a complete picture of SMS (Delignières & Torre, 2011; Wing & Beek, 2002). Some believe the two approaches are entirely reconcilable in the form of a unified theory (Pressing, 1998b, 1999a), while others point out the futility of contrasting two theories that are largely based on divergent methodologies (Balasubramaniam et al., 2004; Torre & Balasubramaniam, 2009).

Another possibility that has been overlooked in many studies is that emergent and event-based timing modes may work cooperatively in separate parts of the body (MacRitchie, Varlet, & Keller, 2017). For example, while a musician is executing discrete, event-based actions with their fingers, they may also exhibit temporally structured emergent movements in their upper-body. This would provide a continuous signal of timing that might indicate to co-performers an individual's personal sense of time, while also reinforcing the pulse intrapersonally. Such a system would be beneficial, as a primary goal in ensemble performance is to minimize asynchronies among co-performers, which is accomplished by anticipating upcoming temporal intervals in order to accurately execute timed actions. Again, given the goal-directed nature of ensemble synchronisation, the underlying mechanism controlling rhythmic production is likely explained to some extent by cognitive accounts of timing. However, musicians and music-listeners alike often exhibit rhythmic body sway that corresponds to the musical pulse, or some other metrical level of the musical phrase structure (Burger et al., 2014). These movement trajectories might initially be self-organized

dynamical systems, but may become very clearly pulse-based at certain points in a performance, allowing musicians to continuously display their sense of time as necessary. Thus, music-related movement and timing might be a matter of constant shifting along a continuum between event-based, cognitively controlled timing, and emergent, dynamical timing. In such a process, musicians entrain in an oscillatory manner within metrical cycles through self-organized ancillary movements, while correcting and anticipating note onsets between cycles through internal representations of discretely timed events (see **Figure 4**). Crucially, this process would depend on co-performers being able to see each other or some other visual rhythm such as a conductor.



Figure 4. A depiction of how timing mechanisms between- and within-cycles may work together in SMS with an external pulse.

# 4. Previous studies of visual cues, ancillary movement, and synchronisation in musical contexts

Investigations of ensemble synchrony should consider visual as well as auditory information; although auditory feedback from oneself and co-performers is sometimes sufficient for synchronising well, concurrent visual cues likely reduce asynchronies further

and facilitate synchrony during musical phrases of irregular tempo. In most large ensembles, a conductor initializes the tempo, providing a visual beat to which the ensemble synchronises (occasionally at a lag if the conductor is maintaining a beat ahead of the ensemble). Furthermore, performers engage in ancillary motions, which can be expressive to convey subjective emotional information (Davidson, 2012; Demos et al., 2014; Thompson & Luck, 2011) but also contain temporal information related to pulse timing (Goebl & Palmer, 2009; Keller & Appel, 2010), tempo changes (Thompson & Luck, 2011), and phrase boundaries (MacRitchie, Buck, & Bailey, 2013).

With a few exceptions, medium- to large-sized ensembles require (or certainly benefit from) a conductor directing rehearsals and performances. The specific roles of a conductor vary according to differences in conducting style and the needs of the ensemble, but nearly all conductors will, among other tasks, dictate the tempo (pulse rate) and tempo changes throughout most of a piece of music (Fredrickson, 1994; Labuta, 1982). Thus, conductors visually relay temporal information to the musicians, who then factor this visual beat into their estimation of time intervals. Given that auditory is generally superior to visual information for perceiving a pulse (Grahn, 2012; Repp, 2003b), it is important to consider what aspects of a conductor's kinematics (and thereby their visual information) are most important for clearly defining the pulse.

Although conductor kinematics have not been experimentally studied extensively, a few studies have found consistent results in both ecologically valid and controlled laboratory settings. One study used point-light models of conductors that were recorded using motion capture and found that quantitatively averaged motion profiles compared to individual original motion profiles resulted in better synchrony for musicians tapping along (Wöllner, Deconinck, Parkinson, Hove, & Keller, 2012), suggesting that musicians are sensitive to subtle differences in conductor motion and work better with smooth prototypical profiles

where signal-to-noise is increased. Recordings of an orchestra rehearsal found that points of maximal synchrony in the ensemble correlated with points of highest deceleration along the movement trajectory of the conductor baton (Luck & Toiviainen, 2006). In other words, a sense of pulse seemed to occur just *after* the point of highest velocity of a vertical conductor gesture. In a tapping task in which participants synchronised with a controllable virtual conductor, the shape of the movement trajectory (measured as radius of curvature) did not predict synchronisation accuracy when velocity was held constant, but velocity (and consequently acceleration) was a significant predictor of accuracy when shape was held constant (Luck & Sloboda, 2008). In other words, high rates of deceleration along a movement trajectory were again related to synchronous behaviour.

Similarly, in a study of the influence of velocity profile on visuomotor entrainment, participants synchronised continuous movements best when the stimulus (a horizontally oscillating dot) accelerated away from a turnaround point, and decelerated when approaching a turnaround point (Varlet, Coey, et al., 2014), a velocity profile that is consistent with biological movements. These studies suggest that people synchronise movements best with visual cues of especially high acceleration, particularly when there is an accentuated slowness in the trajectory toward a turnaround point. In the case of a conductor, this finding could be a rapid increase in the speed of the conductor's trajectory alerts performers to the forthcoming pulse. Thus, the acceleration within a cycle allows one to establish a prediction, thereby facilitating anticipatory timing.

Visual information may also come from co-performers in an ensemble setting via their somatic cues such as body or head sway (Goebl & Palmer, 2009) and apparent digit movements (Goebl & Palmer, 2009; Hove & Keller, 2010). For instance, Keller and Appel (2010) showed that when co-performers could not see each other, they tended to increase body sway amplitude, as if to make their movements and associated sounds more predictable,
acting a *coordination smoother* (Vesper, Butterfill, Knoblich, & Sebanz, 2010) to facilitate synchrony. In an ensemble, such augmented body sway range could help establish a common pulse. Similarly, piano duos exhibited increased synchronisation of head movement under conditions of restricted auditory feedback (Goebl & Palmer, 2009), again showing the importance of somatic cues which may be interpreted as visual cues in interpersonal music production. Furthermore, a study measuring patterns of gazing behaviour between co-performers found that musicians reliably used mutual gazing just before notated tempo changes to maintain synchrony (Kawase, 2014), suggesting that the visual aspects of music performance are important, and of high priority to ensemble members (King & Ginsborg, 2011; Williamon & Davidson, 2002).

Relative to the role of conducting, the role of body sway in conveying specific temporal information are not as well studied. This is understandable given that part of a conductor's job is to establish a pulse and guide the ensemble through tempo changes (Fredrickson, 1994; Labuta, 1982), whereas body sway is often associated with expressivity in music (Thompson & Luck, 2011). However, based on the studies of visuo-motor entrainment described previously, one would expect musicians to entrain their ancillary movements given that they can see each other. One could argue that apparent entrainment is due to a common stimulus such as the auditory pulse, and not due to visual cues. This is a valid consideration, and is addressed in the series of experiments forming this thesis.

The emergence of synchronised ancillary movements among co-performers would be especially useful in a medium to large ensemble, in which members towards the back have an obscured view of the conductor (Wöllner & Auhagen, 2008) and may need to rely instead on the cues of musicians seated in front of them. Indeed, there is an established practice of following section leaders in an orchestra by monitoring their movements (Bishop & Goebl, 2018a; Wöllner & Canal-Bruland, 2010). There is also evidence of musicians relying on

interpersonal visual cues rather than conductor cues, as in a recent case study of serial dependencies in a string quartet (D'Ausilio et al., 2012). Lastly, as mentioned previously, ancillary movements might also be temporally structured to facilitate *intra*personal timekeeping, as vestibular activity induced by swaying could contribute to meter perception (Phillips-Silver & Trainor, 2005, 2007, 2008; Todd & Lee, 2015).

Overall, there are examples of both emergent and event-based timing mechanisms (Wing, Endo, Yates, et al., 2014) in ensemble performance. Ancillary movements may emerge, but can then be used to establish phase-locking among musicians, or to reinforce a sense of one's own timing. Meanwhile, the discrete rhythmic production of note onsets is perhaps better explained by an event-based mode of timing. This synchronisation process, particularly its anticipatory components, could be informed by visual cues from ancillary movements and conductor patterns.

#### 5. Overview of Experiments

The series of experiments presented here examined the influence of visual cues on musical synchronisation, where synchronisation is considered generally as both intentional discrete movements (drum strokes), and emergent ancillary movements. Each experiment comprised three main questions: 1) How does a visual stimulus affect intentional, discrete timing; 2) How does a visual stimulus affect ancillary movements; and 3) How are intentional and emergent timing related (if at all)? Each experiment focused on a different type of visual stimulus. Given that basic music synchronisation and beat-keeping are common abilities, a method of testing musical timing in the general population was developed for this project. This is primarily because the present research is motivated by the widely observed psychological phenomenon of *sensorimotor synchronisation* rather than highly trained musicianship. Also, previous studies on musical timing have, understandably, been run with trained musicians but basic musical synchronisation is achievable by

nonmusicians as well. As such, one question of this thesis is the extent to which musical timing and associated ancillary movements are generalisable behaviours without specific musical training.

Experiment 1 looked at how conductor kinematics affect drumming synchronisation, predictive timing, and ancillary movements in a synchronisation drumming task similar to traditional SMS tapping tasks. Drumming tasks were used instead of tapping for all experiments as people tend to miss fewer beats when drumming compared to finger-tapping (Madison, Karampela, Ullén, & Holm, 2013). Conductor patterns that adhere to velocity profiles previously found to be related to superior synchronisation were expected to improve synchrony, and anticipation of pulse intervals during tempo changes. This was compared to a visual metronome with non-biological motion (no acceleration), and a non-moving image.

Experiment 2 used a similar task and measurements, but in a dyadic context. Pairs of participants drummed together while observing the same conductor as in Experiment 1. However, their visual access to each other and to the conductor was manipulated to study the combined influence of a live co-performer and a conductor.

Experiment 3 again used the same task and similar measures, but a point-light model of a person was used as the visual stimulus. The point-light model is meant to simulate a coperformer whose body sway might provide useful timing information to the participant. This investigated how biological motion that directly resembles the participants' movements can influence synchronisation and ancillary movements.

## **EXPERIMENT 1**

A research report as published in Acta Psychologica, 2018

The influence of visual cues on temporal anticipation and movement synchronisation with musical sequences

Ian D. Colley<sup>a</sup>

Manuel Varlet<sup>a</sup>

Jennifer MacRitchie<sup>a</sup>

Peter E. Keller<sup>a</sup>

<sup>a</sup>The MARCS Institute for Brain, Behaviour and Development, Western Sydney

University, Sydney NSW, Australia

Correspondence concerning this article should be addressed to Ian D. Colley, The

MARCS Institute for Brain, Behaviour and Development, Western Sydney

University, Locked Bag 1797, Penrith, NSW 2751,

Australia. +61 2 9772 6585. Email i.colley@westernsydney.edu.au

Declarations of interest: None

#### 1. Abstract

Music presents a complex case of movement timing, as one to several dozen musicians coordinate their actions at short time-scales. This process is often directed by a conductor who provides a visual beat and guides the ensemble through tempo changes. The current experiment tested the ways in which audio-motor coordination is influenced by visual cues from a conductor's gestures, and how this influence might manifest in two ways: movements used to produce sound related to the music, and movements of the upper-body that do not directly affect sound output. We designed a virtual conductor that was derived from morphed motion capture recordings of human conductors. Two groups of participants (29 musicians and 28 nonmusicians, to test the generalizability of visuo-motor synchronisation to non-experts) were shown the virtual conductor, a simple visual metronome, or a stationary circle while completing a drumming task that required synchronisation with tempo-changing musical sequences. We measured asynchronies and temporal anticipation in the drumming task, as well as participants' upper-body movement using motion capture. Drumming results suggest the conductor generally improves synchronisation by facilitating anticipation of tempo changes in the music. Motion capture results showed that the conductor visual cue elicited more structured head movements than the other two visual cues for nonmusicians only. Multiple regression analysis showed that the nonmusicians with less rigid movement and high anticipation had lower asynchronies. Thus, the visual cues provided by a conductor might serve to facilitate temporal anticipation and more synchronous movement in the general population, but might also cause rigid ancillary movements in some non-experts.

*Keywords:* Sensorimotor synchronisation, temporal prediction, detrended fluctuation analysis, visuo-motor coordination

#### 1. Introduction

Integrating movement and sensory input to interact with the environment with high temporal precision is a fundamental aspect of human behaviour. Such precision is exemplified in music performance, where it is aided by temporal structuring principles that include rhythms consisting of ratio-related durations and hierarchical metrical frameworks (London, 2012). Underlying such frameworks is a subjective sense of regularity known as the *beat* or *pulse* (Iversen & Patel, 2008; Large, Herrera, & Velasco, 2015; Merchant, Grahn, Trainor, Rohrmeier, & Fitch, 2015), which is useful for establishing a shared sense of musical time among people. Auditory-motor connections enable most people to move in time to the beat (Phillips-Silver et al., 2011; Sowinski & Dalla Bella, 2013), whether in the form of rhythmic tapping (Wing, 2002), dancing (Burger et al., 2014), or playing an instrument (Maes et al., 2015). In the latter two cases, timing is frequently coordinated not just intrapersonally by an individual attempting to keep a steady beat, but also interpersonally, as seen in ensemble performance (Keller, 2008; Rasch, 1979).

In a musical ensemble, several, sometimes dozens of musicians aim to coordinate their actions to produce the desired sound within a small window of temporal precision. While the presence of a beat is useful for synchronisation, beat-based ensemble music rarely features a single, repeated beat interval (i.e. *isochrony*). Through expressive interpretation (Repp, 1998; Thompson & Luck, 2011) and notated tempo changes (Loehr, Large, & Palmer, 2011; Repp & Keller, 2004; van der Steen, Jacoby, Fairhurst, & Keller, 2015) the beat rate fluctuates. This leads to the general topic of our investigation: the ability of individuals to synchronise their actions with exogenous tempo-changing rhythmic signals.

Sensorimotor synchronisation (SMS)—that is, the coordination of movements with rhythmic external events—is generally facilitated through perceptual monitoring and reactive error correction (Wing, Endo, Bradbury, & Vorberg, 2014; Wing, Endo, Yates, et al., 2014).

In musical ensembles, this entails listening to oneself and others and responding to interpersonal timing discrepancies. Research on how people keep time has found negative serial dependencies in tap intervals, suggesting people correct successive timing errors. These errors are due to motor system noise, and noise in an internal timekeeper (Wing & Kristofferson, 1973) suggesting that repetitively timed actions are triggered by an imperfect but adaptive cognitive control system (Torre & Balasubramaniam, 2009; Zelaznik, Binsted, Georgescu, & Brownell, 2007). An additional strategy for synchronizing with tempo changes involves prediction, specifically *temporal anticipation* (Mills, van der Steen, Schultz, & Keller, 2015; Pecenka & Keller, 2009a). If musicians can anticipate an upcoming beat interval, then they can minimize the error that they will need to correct, and thus achieve a more cohesive ensemble sound. Both *correction* and *anticipation* have been considered in a more recent timekeeper model (van der Steen & Keller, 2013), which can account for tempo changes by adapting to fluctuating time intervals.

From numerous SMS experiments (Repp & Su, 2013)—in which participants tap a finger in time to a pacing signal or other auditory stimulus—we know that there is a tendency for individuals to predict upcoming time intervals. Although people vary in their ability to predict (Colley et al., 2017; Mills et al., 2015), temporal anticipation is not a static skill. Instead, it can be improved by partnering with another individual who is a good predictor (Pecenka & Keller, 2011), and by observing visual cues that are informative about event timing (Maruta, Heaton, Kryskow, Maule, & Ghajar, 2013; Repp & Moseley, 2012). These strategies—partnering and visual cues—are relevant in musical ensemble performance, especially in large ensembles (e.g., symphony orchestras) where individuals rarely play alone but rather in a section (e.g., a group of violinists all playing the same part), and musicians can see each other as well as a conductor. The present study will focus on the role of visual cues

such as those provided by the conductor, and how these cues influence musical synchronisation.

The conductor is typically responsible for interpreting expressive aspects of the music, and also directing the musical timing. Thus, he or she provides a temporally relevant visual cue in the form a gesture made with a handheld baton. Musical beat locations are usually marked by a rapid downward trajectory of the baton, thus providing a common source of continuous visual information to the ensemble musicians (see **Figure 2.1C**).



Figure 2.1. Diagrams of the visual cue conditions. Panel A depicts the stationary circle, which did not move during trials. Panel B depicts the no-acceleration cue, which reached its lowest point at beats 1 and 3, and its highest point at beats 2 and 4. Panel C depicts a common conductor gesture, which was performed by our conductors when making the virtual conductor. Each number correspond to a musical beat, and after every fourth beat, the gesture repeats.

Fundamental research on visuo-motor coordination suggests that marked deceleration towards the endpoint of a moving object's trajectory makes the timing of the endpoint more salient, thereby facilitating synchronisation with that object (Varlet, Coey, et al., 2014; Zelic, Varlet, Kim, & Davis, 2016). Consistent with this, studies on synchronising with conducting gestures have shown advantages for specific types of motion trajectory. For example, musicians were able to synchronise better with a "morphed" virtual conductor that was made by averaging the movements of multiple conductors, then with individual conductors (Wöllner et al., 2012). Presumably, this is because the morphed conductor provided a prototypical gesture with minimal noise (i.e., minimal variability along the trajectory), so the target timing could be readily predicted. In the same study, the reported quality of the virtual conductors correlated with the vertical velocity of the gestures, suggesting that people are sensitive to subtle differences in visual cues, and that fast vertical motion between beat locations is important for conveying time. Similarly, a study of an ensemble rehearsal found that moments of maximal synchrony in the orchestra correlated with the vertical velocity of the conductor's baton (Luck & Nte, 2008; Luck & Toiviainen, 2006).

However, none of these studies have looked explicitly at whether such visual cues specifically affect the process of temporal anticipation by quantifying the degree to which movements are aligned with tempo-changing beat intervals. This has been considered outside of the music domain in a study that had participants anticipate the action timing of a human or robot (Saygin & Stadler, 2012), but the kinematics were kept constant and only the appearance of the stimulus changed. Given the role of kinematics in synchronisation (D'Ausilio et al., 2012; Luck & Nte, 2008; Varlet, Coey, et al., 2014) and the apparent effects of visual cues on anticipation (Knoblich & Flach, 2001; Koul, Cavallo, Ansuini, & Becchio, 2016; Schubotz, 2007; Wöllner & Canal-Bruland, 2010), we investigated whether continuous visual information with a dynamic velocity profile can facilitate predictions of upcoming beat

intervals, thereby improving synchronisation performance. Furthermore, these beneficial effects of visual information could arise directly by influencing estimates of the timing of upcoming sounds or indirectly by entraining the body movements of the individual.

Studies on visuo-motor entrainment in the field of ecological psychology have shown that individuals entrain their movements to visual rhythms in the environment, sometimes even unintentionally (Richardson et al., 2007; Schmidt et al., 2007; Varlet, Bucci, et al., 2015). In studies of body motion during music performance, a distinction has been drawn between instrumental movements, which are directly related to the production of musical sounds (e.g., the keystrokes of pianists), and ancillary movements, which are not causally linked to sound production (e.g., head nods or body sway) (Nusseck & Wanderley, 2009). The functions of ancillary movements may be related to expressive aspects of the music (Castellano, Mortillaro, Camurri, Volpe, & Scherer, 2008) and also to the control and communication of performance timing (Ginsborg & King, 2009; Goebl & Palmer, 2013).

In line with their functional distinction, instrumental and ancillary movements may be linked to different levels of musical structure. Music typically has hierarchical time scale structure, such that rapid events unfolding at short time scales are embedded within slower, less frequent events at longer time scales. Instrumental movements often account for the fastest events (beats, or beat divisions), and ancillary movements may relate to larger time scales, such as bars formed from multiple beats, or phrases formed from multiple bars (MacRitchie et al., 2013; Thompson & Luck, 2011).

Ancillary movements may also be communicative, conveying timing cues to coperformers when auditory information is degraded or reduced (Goebl & Palmer, 2009). Consistent with communicative functions, it has been found that co-performers make more eye contact during irregularly timed musical passages, suggesting that visual information is especially beneficial when interpersonal coordination demands are high (Kawase, 2014).

Thus, in addition to the conductor, musicians rely to some degree on seeing each other to maintain a shared sense of time, and may capitalize on this by moving rhythmically. Furthermore, movement kinematics in orchestral musicians were found to relate to leadership in the orchestra (D'Ausilio et al., 2012), again pointing towards the importance of body movement in musical synchronisation.

While the majority of studies on sensorimotor synchronisation have focused on sound-producing instrumental movements, there is evidence that ancillary movements can also play an important role in time-keeping. Performers may use ancillary movements to act as a "coordination smoother" (Vesper et al., 2010) to make their actions more regular, and thus predictable, as seen in a study where pianists in duos increased the amplitude of their body sway when they could not see each other (Keller & Appel, 2010). The effects of increased body sway on timing regularity may be related to increased head movement, which may facilitate timing by reinforcing one's sense of rhythm through the stimulation of vestibular networks (Phillips-Silver & Trainor, 2008; Todd & Lee, 2015; Trainor, Gao, Lei, Lehtovaara, & Harris, 2009).

However, despite evidence that whole-body movements may be related to timekeeping, it is not conventional to measure ancillary movements during basic SMS tasks, and those that examine ancillary movements in skilled music performance (Goebl & Palmer, 2009) necessarily exclude nonmusicians, despite the fact that musical synchronisation is a widespread phenomenon (e.g. audience members tap or nod with the music). The aim of the present study was to investigate the effects of continuous information in visual cues provided by conducting gestures on synchronisation with musical sequences containing tempo changes. We assumed that in addition to influencing instrumental movement by improving temporal anticipation of musical beat locations (relative to no visual stimulus, or a simple moving stimulus), continuous visual cues provided by rhythmic conducting gestures might

also entrain rhythmic ancillary movements. This could in turn enhance an individual's SMS ability by improving the stability of one's embodied sense of time.

The task used to assess the role of visual cues in motor coordination was an SMS tapping task. We tested both highly trained ensemble musicians, and people with no formal musical training, to test the generalizability of visual cues in music synchronisation beyond people with relevant experience. Typical synchronisation studies use click tracks devoid of pitch variation and harmony (Keller & Repp, 2008; Repp, 2008; Zelaznik et al., 2005). This is effective for studying timing outside the music domain, but we were specifically interested in musical timing. Therefore, our stimuli were designed with harmonies and multiple instruments to simulate a musical context with tempo changes, in which a conductor is typically considered to be helpful. Previous studies have found that individual differences in auditory imagery predict anticipatory timing abilities, and that people synchronise more accurately with isochronous (or nearly isochronous) pacing signals and music excerpts than with tempo changing pacing signals (Pecenka & Keller, 2009a) or expressively timed music excerpts (Colley et al., 2017). Knowing this, we wanted to focus on how anticipatory timing can be influenced or improved across long musical sequences with alternations between steady phases and tempo-change phases. Therefore, participants were instructed to drum in synchrony with the beat of rhythmically simple, but tempo-changing, multi-part music, for 3 m 22 s per trial. Concurrently, in separate conditions, participants observed a virtual conductor, a simple visual metronome that moved without acceleration (within each cycle) on the vertical axis only, or a stationary circle (essentially a hearing-only condition). We hypothesized that the virtual conductor would reduce asynchronies and improve temporal anticipation relative to both the non-moving stimulus and the simple visual metronome, due to the virtual conductor's informative velocity profile (Luck & Toiviainen, 2006; Varlet, Coey, et al., 2014), which clearly marks beat locations. We also expected the degree of

temporal anticipation to correlate negatively with asynchrony, as increase in the ability to predict the timing of beats due to the conductor should supplement the general tendency for greater prediction to result in lower asynchronies (Colley et al., 2017). Furthermore, we expected these effects to apply to both experts (ensemble musicians) and nonexperts (nonmusicians) as the underlying kinematic processes of visuo-motor and audio-motor synchronisation can be observed in humans regardless of musical experience (Colley et al., 2017; Hove & Keller, 2010; Zelic et al., 2016).

To examine ancillary movements, we measured the upper-body and head motion of participants. A commonly applied method of quantifying spatiotemporal structure in body movement, including body sway (Blázquez et al., 2009), is detrended fluctuation analysis (DFA), which can categorize time series by noise colour (Kantelhardt, Koscielny-Bunde, Rego, Havlin, & Bunde, 2001; Peng et al., 1994). This measure gives an indication of how structured movements are, on a scale from random (un-structured white noise) to deterministic (very structured Brownian noise); the measure does not give an indication of the shape of movement, and the shape of movement patterns was not a focus of this project so much as the timing and regularity of movement cycles. We hypothesized that head movements would be more structured (i.e. follow a pattern of motion) and of greater magnitude in the virtual conductor condition, as the naturalistic visual rhythm was expected to entrain ancillary motion (Schmidt et al., 2007). As a means of linking instrumental and ancillary movements, we also expected a negative relationship between body movement structure and asynchrony, meaning those who move in a more structured way would show lower asynchronies (Phillips-Silver & Trainor, 2008).

## 2. Methods

2.1. Participants. We recruited 29 musicians (14 female) aged 18-50 years, and 28 nonmusicians (19 female) aged 18-35 years. Criteria for inclusion as a musician were five or

more years of musical training (median = 14.43 years, range = 10-37 years), currently practicing/performing, and, having had experience playing with a conductor. The criterion for nonmusicians was having no musical training. This rather strict criterion was upheld by listing it in the study advertisements, and verifying with participants when they signed up for the study, and veryifying when they arrived. Despite having no musical training, all nonmusician participants reported listening to music on a daily basis. Most participants were recruited through the Western Sydney University School of Psychology, and received course credit. Some of the musician sample were recruited from various music ensembles around Sydney. They were paid \$20 to reimburse travel costs. All participants provided written informed consent prior to the experiment, which was approved by the Western Sydney University Ethics Committee.

2.2. Study Design. The experiment was a  $2\times3$  mixed design. There were two *expertise* levels as the between-subjects factor (musician and nonmusician) and three levels of *visual cue* as the within-subjects factor (stationary circle, no-acceleration motion, and conductor). The dependent variables were mean absolute asynchrony, an index of temporal anticipation, standard deviation of movement, and fluctuation of movement (i.e.,  $\alpha$ DFA).

*2.3. Apparatus.* An Alesis Percpad (tapping pad) was used to collect the synchronisation data. Participants used a drum stick rather than finger, as recent studies have shown that synchronisation drumming results in fewer missed taps than synchronisation tapping (Madison et al., 2013; Manning, Harris, & Schutz, 2017). Participants' movements were recorded with a 12-camera Vicon motion capture system at 100 Hz sampling rate, with reflective markers arranged using the built-in upper body model in the Nexus software package. The motion capture recording and the drum recording were synced by sending a serial trigger signal to Nexus at the onset of each trial. The experimental procedure (data collection, stimuli presentation, and trigger signals) was programmed using the

OpenFrameworks coding environment for C++ on a 2015 MacBook Pro. Auditory stimuli were sent through stereo speakers, and visual stimuli were presented on a 17" monitor with a 60 Hz refresh rate.

2.4. Auditory stimuli. Three short pieces (3 m 22 s) were created (by author IDC) for the experiment using the notation software *Musescore*. The intention was to create stimuli with a constant and unambiguous beat, but with some melodic and harmonic interest to simulate a musical setting. Thus, the only rhythmic values used were quarter notes in the upper two voices (glockenspiel and xylophone) and eighth notes in the lower voice (harp). These instruments were chosen as they had rapid onsets and were voted as the most pleasing MIDI instruments during pilot testing. There were no rests (i.e., silent beats), meaning every beat as defined in 4/4 meter included an audible note in the music. The pitch range was C2-A5, which is well within typical musical ranges. Melodies and harmonies were based on basic practices in Western music theory. The xylophone and glockenspiel played complementary melodies, while the harp accompanied with chords. Full scores are included in **Appendix A**. The length was chosen to reflect a typical short piece of music and to allow for more reliable analyses of motion capture data, as discussed later.

To create the tempo changes, the Musescore files were exported as MIDI files, which were converted to ASCII format, then edited in Matlab to change the note on/off times. The music started at 120 bpm or a 500 ms inter-onset interval (IOI). This steady tempo phase continued for eight beats, then a tempo change would occur over eight beats, either slowing or accelerating. The direction of change would then reverse to bring the music back to 120 bpm for eight beats. Thus, the location of tempo changes was regular, but the direction of change, and the magnitude of change were randomly generated. Each participant heard the same music with the same tempo changes. There were six rates of change for the tempo changes:  $\pm/-10$ ,  $\pm/-16$ , and  $\pm/-22$  ms per beat. These rates of change were chosen based on

pilot test results. After editing in Matlab, the files were saved as MIDI files, then opened in Garageband to set the instruments for each track, and lastly saved in AIFF format.

2.5. Visual stimuli. There were three visual stimuli: the stationary circle (Figure 2.1A), the no-acceleration circle (Figure 2.1B), and the virtual conductor (Figure 2.1C). All three stimuli used a red circle with 13 mm diameter against a black background. These colours were chosen based on a participant suggestion during pilot testing, and the remaining pilot participants preferred red on a black background over black on white. The no-acceleration circle moved vertically between two turnaround points 132 mm apart, with a constant speed within each cycle. Turnaround points always corresponded to a musical beat (e.g. beat one corresponded to the lowest position, beat two to the highest). Thus, the speed between cycles would change instantaneously to match the tempo changes, but the speed within a cycle would not change, hence the nomenclature "no-acceleration." We made the virtual conductor by averaging the motion capture recordings of three conductors (Wöllner et al., 2012). The resulting position coordinates determined the trajectory of the circle during a conductor trial. In addition to changing speed within a beat cycle, the conductor differed from the no-acceleration stimulus by moving horizontally as well as vertically, as is common for a conductor pattern.

2.6. Virtual conductor. The three conductors contacted to participate in motion capture recording as part of the design for the virtual, morphed conductor had at least 10 years experience conducting a variety of ensembles, including student string orchestras, full symphonic orchestras, and an army band. They were reimbursed \$30 for their assistance. We recorded and then averaged their distinct conducting styles in order to arrive at a prototypical average that reduces noise from individual conductors (Wöllner et al, 2012) thereby making the stimulus more generalisable.

The conductors were sent the audio files and scores of the music one week before the scheduled motion capture recording. They were asked to practice conducting the three pieces at least twice per day (about 20 minutes per day), or until they "felt as if [they] were leading a small ensemble," and encouraged to notate the scores in whatever way was useful. On the day of the recording, the conductors—who were recorded individually—were outfitted with reflective markers according to the upper-body model included in the Nexus software (a motion capture program that recorded the motion of markers, which were attached to conductors with tape and elastic bands). However, the model was edited to include a baton with two additional markers: one at the handle, and one at the tip. Although we intended to only use the baton tip marker for this experiment, we recorded the full upper-body for use in future studies.

We recorded three takes of each of the three pieces for a total of nine takes. The experimenter started the music, which was preceded by four count-in beats using a woodblock sound. The conductors were asked to start conducting on the third count-in beat in a 4/4 pattern (Figure 2.1). A trigger signal was sent to the motion capture system when the music count-in started so the recordings could be synced offline.

To average the motion capture recordings, we exported the position data of the baton tip as an ASCII file. The frames were trimmed to start at the trigger signal and end one second after the final beat. To match the motion capture to the refresh rate, the recordings were down-sampled from 100 Hz to 60 Hz, and filtered with a 10 Hz low-pass Butterworth filter in Matlab. The resulting vectors were 3D position coordinates of the baton tip, but we only used the *x* and *z* planes to make the 2D virtual conductor (the *z* plane in the motion capture software corresponds to the vertical plane, or *y*, in 2D Cartesian space, which would be the coordinate system for the stimuli). The vectors were shifted so the minimum value was zero, and scaled to fit within the computer screen. Lastly, the vectors were averaged using a simple arithmetic mean (as in Wöllner et al., 2012), and saved as text files.

2.7. Procedure. Upon arrival, participants were briefed on the task, then read and signed a consent form, and filled out a questionnaire of musical experience to verify that they met the criteria for either musician or nonmusician. Next, the experimenter attached the reflective markers for the motion capture recording according to the built-in upper-body model in Nexus. Participants were contacted before the day of testing and asked to wear a tight-fitting shirt if possible, so as to minimize extraneous motion of the markers. As the experimenter attached the markers, he explained the task. Participants were instructed to stand on a marked location in front of the testing monitor, which was placed on a high table and adjusted so the centre of the screen was at eye level. They were told to stand comfortably and that they were free to move, so long as they continued to face the monitor. The experimenter demonstrated how to hold the drum stick for all participants, regardless of musical experience. All participants were able to follow this demonstration. No other explicit instructions regarding movement were given.

For each trial, the experimenter would start the motion capture recording, then prompt the participant to start the music by pressing the 'return' key on the testing computer when they were ready. This would begin the four-beat count-in, and participants would start drumming with the music after the fourth count-in beat. As expected, some nonmusician participants were unfamiliar with the concept of a count-in, and so the four-beat count-in was explained and demonstrated during a practice trial. There was no electronically generated auditory feedback from the drum, just the sound the stick hitting the drum. To ensure they were observing the visual stimuli, a letter would appear at random points throughout a trial in the middle of the screen, and participants were told to say the letter out loud. This is similar to procedures in other visuo-motor synchronisation studies (Varlet, Bucci, et al., 2015;

Varlet, Coey, et al., 2012). Due to the length of our trials, several letters would appear throughout every trial to sustain participants' attention. There was one practice trial, which used the stationary circle. Then, each of the three pieces of music was paired with the three visual stimuli twice  $(3 \times 3 \times 2)$  making 18 experimental trials. This was divided into three blocks of six trials to give participants breaks. Each block contained the three visual conditions twice, but in a random order. Participants were instructed to keep in time with the music even as it changed speed, to continue drumming until the music stopped, and to always watch the monitor. After the experiment, participants were debriefed and asked for feedback about the experiment and usefulness of the visual cues.

#### 3. Data Analysis

3.1. Drumming SMS. There were two main dependent measures of instrumental movements (i.e. the drumming data) which come from the theory of internal timers from the school of information processing: mean absolute asynchrony and an index of temporal anticipation. For both measures, the inter-drum interval (henceforth inter-tap interval [ITI] for consistency with other timing studies) series needed to be the same length as the IOI series of the music. On average, 83% of the trials in the musician group, and 67% of trials in the nonmusicians group had an equal number of ITIs and IOIs. If the series lengths did not match, we used the following interpolation procedure: ITI values that were twice as large (with a tolerance of +/- 10%) as the corresponding IOI were split into two equal values to account for the presumed missed tap. The same was done for ITI values that were three and four times as large as the corresponding IOI, but split into three and four equal values respectively. Any trials with more than three consecutive missed beats were discarded. If an ITI was less than 100 ms, it was considered a double-tap (meaning two successive and rapid taps occurred in the space of one musical beat), and added to the previous ITI under the assumption that the sum of the two successive intervals represent the participant's intended

tap interval. Once the series were the same length, we calculated the mean absolute asynchrony as a general representation of how far off from the beat participants were on average. To do this, we subtracted the cumulative IOI from the corresponding cumulative ITI at each beat, converted the differences to absolute values, and averaged this asynchrony series. We removed trials for which the mean absolute asynchrony was 500 ms or greater, as this was the average IOI in the auditory stimuli. This was about 3% of all trials across all participants.

Temporal anticipation was quantified using a *prediction/tracking index* (P/T index) using cross-correlation (CC; Colley et al., 2017). If participants are anticipating IOIs, then the ITI series should resemble the IOI series at lag-0 (i.e. the actual IOI series). If they are tracking the tempo changes, then the ITI series will resemble lag-1 of the IOI series. By dividing the coefficient of the lag-0 CC by that of the lag-1 CC, we get a measure of the extent to which individuals are predicting (quotient > 1) or tracking (quotient < 1). The main analysis was a  $2 \times 3$  mixed ANOVA with *expertise* as the between-subjects factor (musician or nonmusicians), and *visual cue* as the within-subjects factor (stationary circle, no-acceleration, conductor). This was done for both mean absolute asynchrony and the P/T index.

*3.2. Ancillary movements.* Our analysis of ancillary movements comes from analytical methods of the dynamical systems school. We focused on one marker located on the head (the right forehead marker), as we found that most participants moved their head rather than torso during pilot testing (we still recorded the whole upper-body for use as a visual stimulus in future studies). Furthermore, we were specifically interested in using motion capture to understand ancillary movements, so we did not analyse the arm movements, which are considered instrumental movements. To standardize the movement volume of participants, the four markers around the hips were averaged to create a centre point for each trial, which

was used as the origin for the other markers. To reduce processing time of the series, the 100 Hz recordings were down-sampled to 50 Hz. Next, the series were filtered with a 10 Hz lowpass Butterworth filter. To assess the amplitude of movement, we calculated the standard deviation (Stoffregen, Fu-Chen, Varlet, Alcantara, & Benoît, 2013; Varlet, Bardy, Chen, Alcantara, & Stoffregen, 2015; Varlet, Stoffregen, et al., 2014) of the position coordinates on each of the three axes. To assess fluctuations and how structured the movements were, we used DFA in the RStudio package "nonlinearTseries." The primary dependent measure that is given by DFA is the scaling exponent,  $\alpha$ , which ranges from 0.5 (white noise/random behaviour) to 1.5 (Brownian noise /deterministic behaviour). A value of 1.0 indicates pink noise, which is associated with default coordination in movement such as standing balance sway (Blázquez et al., 2009; Wang & Yang, 2012).

DFA works by first breaking a time series into windows of size *n*. The time series within each window is detrended (usually linearly), and analysed for variance. The variance is then averaged across all windows to produce a fluctuation value at that window size. The size of *n* is then increased to the next power of two, and the process repeated until *n* is about half of the whole series length. Alpha ( $\alpha$ DFA) is the slope of the regression relating fluctuation to each window size, and therefore represents how a system operates over multiple time scales. We used window sizes from 2 to 4,096, where 2 is the smallest power of two, and 4,096 is a power of two that is about half our time series length. We used a regression range from window size; window sizes above 1,000 tended to produce exponential increases in variance, which would overestimate  $\alpha$ DFA. We again used 2 × 3 (*expertise* by *visual cue*) mixed ANOVAs, but this time ran separate tests for the three spatial axes (*x*, *y*, *z*, henceforth side-to-side, forward-backward, and up-down, respectively). This was performed

for each of the two motion capture dependent measures (standard deviation of movement and  $\alpha$ DFA).

3.3. Multiple regression. To relate the head movement data to the drum timing data, we used a multiple regression model with  $\alpha$ DFA and P/T index as predictors of asynchrony. Asynchrony was chosen as the dependent variable as it was the measure of performance success; participants were instructed to synchronise, and low asynchrony is desirable in most music performances.  $\alpha$ DFA represents participants' behaviour at the level of ancillary movements (head movements), and P/T index represents a process that we assume is related to instrumental movements (drum timing).

#### 4. Results

4.1. Asynchronies and Prediction/Tracking Indices. For asynchrony (see Figure 2.2) there was a significant main effect of *expertise*,  $F_{(1, 55)} = 9.97$ , p < .01, partial  $\eta^2 = .15$ , such that musicians produced lower asynchronies (were more accurate) than nonmusicians. There was also a main effect of *visual cue*,  $F_{(1.32,72.65)} = 4.21$ , p < .05, partial  $\eta^2 = .07$  (Greenhouse-Geisser corrected), such that the conductor condition produced lower asynchronies than both the stationary circle (p < .001) and the no-acceleration conditions (p < .05), Bonferroni corrected. There was no difference between the stationary circle and no-acceleration circle conditions.

For the P/T index (temporal anticipation; see **Figure 2.3**), there was no main effect of *expertise* or *expertise* by *visual cue* interaction. There was a main effect of *visual cue*,  $F_{(2,110)} = 13.40$ , p < .001, partial  $\eta^2 = .20$ . Bonferroni corrected post-hoc comparisons showed the conductor condition yielded higher P/T indices than both other conditions, p < .001.

There was also a significant negative correlation between asynchrony and P/T index (averaged over visual cues) across all participants,  $r_{(55)} = -.24$ , p < .01, indicating higher P/T index related to lower asynchrony overall. This relation was not driven by one group in

particular, as the correlation was significant for musicians,  $r_{(27)} = -.25$ , p < .05, and for nonmusicians,  $r_{(26)} = -.24$ , p < .05 across all visual cues.



**Figure 2.2**. Mean absolute asynchronies for each expertise group (musician and nonmusician) in each visual cue condition. Error bars represent standard error.



**Figure 2.3**. Mean P/T Index for each expertise group (musician and nonmusician) in each visual cue condition. Error bars represent standard error.

4.2. Motion Capture Analysis. First, we considered the standard deviation of movement, which is a measure of movement amplitude (see Figure 2.4, panels A and B), which was positively skewed for all conditions and groups, so a log-10 transform was used. For the side-to-side movements there was a main effect of *expertise* such that the nonmusicians moved more (higher standard deviation) than the musicians, though the effect size was quite small,  $F_{(1, 55)} = 30.32$ , p < .000001, partial  $\eta^2 = .06$ . However, there was no effect of *visual cue* on standard deviation of side-to-side movement. For the forwardbackward and up-down movements, there was no main effect of group or visual cue.

Next we considered  $\alpha$ DFA (**Figure 2.4**) along each axis of movement. For the sideto-side axis there was no main effect of *expertise* or of *visual cue*. However, there was an interaction effect of *expertise* by *visual cue* along the side-to-side axis,  $F_{(1.46, 80.28)} = 14.37$ , p < .001, partial  $\eta^2 = .21$  (Greenhouse-Geisser corrected). This indicated that for the nonmusician group,  $\alpha$ DFA values were significantly higher in the conductor condition compared to both the no-acceleration (p = .001) and the stationary circle (p = .01), and values in the no-acceleration condition were significantly higher than the stationary circle (p < .05), Bonferroni corrected. The effect of *visual cue* on side-to-side axis  $\alpha$ DFA for musicians was not significant. For both the forward-backward and up-down axes, there were no significant effects of group, visual cue, or interactions.

4.3. Relating movement, prediction, and synchronisation. Lastly, given the effect of visual cue on  $\alpha$ DFA of head movements and on P/T indices for the nonmusician group, we tested whether these two variables ( $\alpha$ DFA and P/T index) could predict asynchrony in a multiple regression. We used values from all three visual cue conditions, but limited the

analysis to the nonmusician group, as this was the only group that showed an effect of visual cue on movement. The regression was significant, F(2, 81) = 5.94, p < .01. Both  $\alpha$ DFA (standardized  $\beta = .22$ , 95% CI [11.0-481.08], p < .05) and P/T index (standardized  $\beta = -.32$ , 95% CI [-1427.94- -294.85], p < .01) were significant predictors of asynchrony with a total R<sup>2</sup> of .12. Nonmusicians' asynchronies were low to the extent that head motion was unstructured and temporal anticipation was high (see Figure 2.5).

## 5. Discussion

This experiment investigated how visual cues that are relevant to music performance affect sensorimotor synchronisation with tempo-changing auditory sequences at both the level of instrumental movements (drumming accuracy) and ancillary movements (head motion). Overall, we found that visual cues that are derived from conductors' gestures can improve temporal anticipation and synchronisation performance both for people with musical experience, and for people with no musical training. This supports our hypothesis that temporal information provided by continuous visual cues containing salient changes in acceleration can improve the prediction of beat timing and thereby facilitate synchronisation with auditory sequences.

More generally, our findings suggest that two multimodal cues that occur simultaneously and provide non-conflicting information can improve SMS relative to a single cue, presumably through efficient multisensory integration (Elliott, Wing, & Welchman, 2014; Ernst & Bülthoff, 2004). The effects of visual cues on ancillary movements appear to be less generalizable, as we found an effect on movement fluctuations for nonmusicians only. This partially supports our hypothesis that visual cues provided by a virtual conductor encourage larger and more structured ancillary movements, but the results also suggest that these effects are modulated by musical expertise. Within the musician group, the conductor stimulus was associated with lower asynchronies and higher P/T indices relative to both other conditions. This means that participants were better able to synchronise their drumming with the tempo changes in the pacing sequence when they observed visual cues based on conductor gestures, presumably because the continuous and naturalistic nature of these cues provided temporal information



**Figure 2.4**. Average standard deviation of movement (log 10 transformed, panels A and B) and average αDFA (panels C and D). Expertise group is separated by panel, and direction of movement is separated by line type within a plot. Error bars represent standard error.

about the onset of upcoming beats. Interestingly, when questioned following the experiment, 11 of the 29 musicians reported that the two moving visual cues were not helpful, or sometimes even distracting. Given that most of the musician sample could synchronise quite easily with just the auditory stimulus (as seen in the stationary circle condition), any difference in performance between conditions may have gone unnoticed by these participants due to a near-ceiling effect. However, as reported above, the conductor did objectively improve performance slightly but significantly compared to the stationary circle and the noacceleration circle. This supports the idea that compatible cues from different modalities (Hills, Ernst, Banks, & Landy, 2002), as well as a dynamic velocity profile with clear deceleration towards a turnaround point (Balasubramaniam et al., 2004; Luck & Toiviainen, 2006; Varlet, Coey, et al., 2014), are beneficial to performance even if skilled participants remain unaware of these benefits. A future experiment might vary the lead/lag time of the virtual conductor relative to the musical beats to test the sensitivity of musicians to a preemptive conductor; that is, a conductor who conducts ahead of the beat.

The nonmusicians also benefited from visual cues, though their asynchrony scores were much higher and more variable compared to the musicians. To validate the relation between asynchrony and P/T index measures, we correlated the two variables within both groups, and across all participants. Indeed there were significant negative correlations as expected, suggesting that high prediction resulted in lower asynchrony. While a few conductor studies have found that synchronisation relates to high acceleration of the

conductor baton (Luck & Nte, 2008; Luck & Toiviainen, 2006), they were not concerned with the mechanism through which acceleration impacts visuo-motor synchrony. Our finding that temporal anticipation was greatest with the conductor visual cue in this study provides causal evidence that, as hypothesized, the way in which a conductor helps an ensemble is, at least partly, by facilitating temporal predictions.



**Figure 2.5.** Scatterplots showing the relationship between movements fluctuations and asynchrony (panel A), and the P/T index and asynchrony (panel B).

For both expertise groups, the no-acceleration stimulus had no effect on drumming performance relative to the stationary circle. This is important when set against the effects observed with the virtual conductor stimulus, as it suggests that visuo-motor timing is influenced by noticeable changes in speed of a moving stimulus, rather than by the motion itself. This interpretation is consistent with previous studies that found that particular velocity profiles are important for visuo-motor tracking both in ecologically valid conducting (Luck & Nte, 2008; Luck & Toiviainen, 2006) and more basic experimental tasks (Varlet, Coey, et al., 2014; Zelic et al., 2016). However, the no-acceleration stimulus in this experiment was not representative of natural human movement, as periodic motion in the human motor system is not constant in velocity. Thus, the no-acceleration condition could be considered incompatible with participants' actions (Hove et al., 2010; Lacquaniti, Terzuolo, & Viviani, 1983; Saygin & Stadler, 2012; Viviani, 2002; Viviani & Flash, 1995)-especially during the circle's upward movement (Hove et al., 2010)-and this could explain why it was not as helpful as the virtual conductor. However, a study of visuomotor synchronisation with similar visual stimuli to our own (Hove et al., 2010), albeit presented at a steady tempo, found a benefit of non-biological stimuli where we did not. This could be because the tempo changes our auditory stimuli were comparatively difficult to synchronise with, and the sounds were the synchronisation target rather than visual signals alone. Again, this lends support for the importance of a dynamic velocity profile in human movement coordination.

The results of the kinematic analysis showed only a small effect of visual cue on ancillary movement, but this still offers some intriguing considerations regarding movement and music. First, the group difference in movement amplitude is interesting as it shows that

for this task, nonmusicians were moving their heads more than musicians. This could be because the nonmusicians found the task more challenging (as evidenced by their larger asynchronies) and needed to embody the beat by activating vestibular networks (see Todd & Lee, 2015) more than musicians needed to, as musicians are generally better able to predict tempo changes and correct errors when synchronising with auditory sequences alone (Manning et al., 2017; Pecenka & Keller, 2011; Repp, 2010). Indeed, there are some studies showing how movement, particularly of the head, can help establish or reinforce musical beat and meter (Phillips-Silver & Trainor, 2005, 2007, 2008), that head movements increase with difficulty in a musical task (Goebl & Palmer, 2009), and that spontaneous movements relate to beat-intervals (Toiviainen, Luck, & Thompson, 2010). The musicians on the other hand may have moved more with syncopated (complex) rhythms, rather than predictable beat sequences that have been shown to be too simple to induce movement (Witek et al., 2017). Ancillary movements might sometimes be communicative (Kawase, 2014; Keller & Appel, 2010) or serve as coordination smoothers (Vesper et al., 2010), leading musicians to move more in the presence of a co-performer, which could be addressed in a future study. However, both groups showed low levels of movement overall, so while the group difference is statistically significant, the effect size is small.

We also found varying responses to visual cues depending on expertise, as seen in the side-to-side movements, which were more structured—meaning the movements followed a more rhythmic pattern—in the conductor condition than the other two conditions for the nonmusicians only. That is, as the complexity of the visual cue increased (no movement  $\rightarrow$  no-acceleration  $\rightarrow$  conductor) so did the determinism of side-to-side movement. In other words, the no-acceleration motion elicited more structured movement than the stationary circle, and the conductor elicited more structured movement than both other conditions in the nonexpert group. However, within this group, higher  $\alpha$ DFA was associated with higher

asynchrony, suggesting that more structured movement might indicate rigidity, which is not conducive to synchronisation.

This finding could be a matter of criticality (Bak et al., 1987), the idea that a dynamical system will stay within a flexible regime (i.e. pink noise, or 1/f fluctuations) in order to be able to adapt to the environment. Musicians, therefore, may have a more adaptable sensorimotor system in this context, which is realized by keeping ancillary fluctuations within a window very close to 1/f. Indeed, this pattern of fluctuation is often associated with a flexible, well-functioning system (Hove et al., 2012; Pressing, 1999b; Wang & Yang, 2012). Some nonmusicians on the other hand, may have transitioned out of a flexible regime into a deterministic or rigid regime. This sort of transition is commonly seen in pathological or sub-optimal performance (Stergiou & Decker, 2011; Wang & Yang, 2012), suggesting that in this task, rigid movements were associated with poorer performance. This finding, while contrary to our hypothesis that more deterministic movements would relate to better performance, corroborates findings that experts tend to show lower alpha values than nonexperts-indicative of greater flexibility-in measures of the experts' domain (e.g. trained runners show lower alphas in measures of running gait cycles compared to nonrunners; Cohen & Sternad, 2009; Nakayama, Kudo, & Ohtsuki, 2010; Wilson, Simpson, Van Emmerik, & Hamill, 2008). While the conductor visual cue was associated with lower asynchronies at the group level, the conductor visual cue also yielded higher aDFA than the other cues. However, aDFA and asynchrony showed a positive relationship in the regression analysis, suggesting that a deterministic regime was prompted by the conductor, but was not helpful for synchronisation for some people.

However, the regression also showed that temporal anticipation (P/T index) can statistically predict lower asynchrony for non-experts. This means that in addition to potential influence from embodied cognitive processes as measured by αDFA, nonmusicians'

synchronisation performance was influenced by anticipatory cognitive processes. With more experience synchronising with tempo changing beat sequences, the nonmusicians might be able to rely on internal predictive processes more, and embodied processes less, or learn to maintain ancillary movements in a flexible range of fluctuations (1/*f*-type fluctuation) as the musicians do.

Future studies might consider the effect of visual cues in synchronisation without an accompanying auditory sequence, in order to approximate the situation faced by ensemble musicians more closely. In this study the auditory stimulus was always presented during a trial, which created a clear objective: match the beat of the music. However, in most cases of music performance, the performers create sound without an auditory beat provided. Thus, it could be the case that participants—especially the musician group—would show more complex or rhythmically salient movements without an auditory sequence, as the task of following a visual reference alone is more difficult (Grahn, 2012; Repp, 2003b). Alternatively, synchronisation without an auditory reference may give participants more control over their actions, allowing ancillary movements to unfold in a more flexible regime. Given evidence that people will use the most reliable modality of timing information (Elliott et al., 2010; Ernst & Banks, 2002), participants might then use proprioceptive cues via ancillary movement if the auditory information is unreliable (i.e. not externally driven). Of course, if the task becomes too difficult movement could become highly irregular or even task-irrelevant. Overall, the role of movement in music is complex, and is likely mediated by numerous variables including individual differences or preferences for movement, the presence of co-performers, the difficulty or rhythmic complexity of the music, and the expressive qualities of the music since movements are not strictly related to timing, but to expression as well (Castellano et al., 2008; Davidson, 2012). Also, we did not consider specific dance or athletic experience in our participants, and these experiences could feasibly

influence one's ancillary movements. As such, our focus on the *general population* limits the conclusions that can be drawn, as the specific motor expertise of individual participants might be important to understanding music-related movement. These issues notwithstanding, the present experiment provides evidence that instrumental and ancillary processes can be modulated by multisensory cues and expertise to influence SMS performance.

## 6. Conclusion

In this study, we have shown that a typical conducting pattern characterized by rapid changes in vertical velocity is, to some extent, causally linked to improved musical synchronisation. The conductor also improved temporal anticipation. This could be taken to suggest that conductor kinematics facilitate the prediction of beat timing, which may possibly be a mechanism through which synchronisation is improved. Interestingly, the effect on prediction was true for both ensemble musicians, and people with no formal musical training. We also examined head movements and found no effect of visual cue on the magnitude or structure of musicians' movements, but a small effect of visual cue on the structure of nonmusicians' movements. These results might suggest that visual cues could be beneficial to interpersonal timing, particularly in the ecologically valid case of a music ensemble. Although the role of ancillary movements in timing is still unclear, we have provided some tentative evidence that conductor gestures promote more structured movements (relative to a simple moving stimulus devoid of acceleration changes) in nonexperts, and that this increase in movement structure was associated with a group-level increase in asynchrony. This suggests that for this task, ancillary movements are most useful in a flexible regime, as seen in the musicians, and that individual differences in ancillary movements can be problematic if they become too rigid as seen in the nonmusicians. However, this process is moderated by experience and context.

# Acknowledgements

Author P.K. is supported by a Future Fellowship grant from the Australian Research Council (FT140101162). We would like to thank conductors David Angell, Marjorie Hystek, and Benjamin North for agreeing to be motion capture recorded for the creation of our stimuli.

## **EXPERIMENT 2**

A manuscript submitted to Human Movement Science

The influence of a conductor and co-performer on auditory-motor synchronisation, temporal prediction and ancillary entrainment in a musical drumming task

Ian Colley

Manuel Varlet

Jennifer MacRitchie

Peter E. Keller

#### Abstract

Ensemble musicians coordinate their actions deliberately in order to achieve temporal synchronisation in their performances. However, there is a tendency for musicians also to move parts of their bodies unintentionally, sometimes in ways that are not directly related to producing sound from their instruments. These movements-intentional or otherwiseprovide visual signals to co-performers, which might facilitate temporal synchronisation. In large ensembles, a conductor also provides a visual cue, which has been shown to enhance synchronisation. In the present study, we tested how visual cues from a co-performer and a conductor affect processes of temporal anticipation, synchronisation, and ancillary movements in a sample of primarily non-musicians. We used a dyadic synchronisation drumming task, in which paired participants drummed to the beat of tempo-changing music. We manipulated visual access between partners and a virtual conductor. Results showed that the conductor improved synchronisation with the music, but synchrony with the music did not improve when partners could see each other. Temporal prediction was improved when partners saw the conductor, but not each other. Ancillary movements of the head were more synchronised between partners when they could see each other, and greater ancillary synchrony at certain frequencies of movement was associated with greater drumming synchrony at certain frequencies of movement. These results suggest that compatible audiovisual cues can improve intentional synchronisation (i.e. drumming with the music). Also, ancillary movements are affected by visual access to a partner, and are related to musical synchronisation under certain visual conditions.
## 1. Introduction

Numerous forms of human behaviour are cooperative, and evoke coordinated movements from the individuals involved. This is seen in social activities such as conversing (Shockley et al., 2007), competitive sports (Varlet & Richardson, 2015), or simply walking together (Nessler & Gilliland, 2009). Even applause of audience members tends to spontaneously synchronise (Néda & Ravasz, 2000). This phenomenon is called unintentional spontaneous entrainment, and refers to the emergence of coordinated or even synchronised behaviour even when a given task does not require explicit coordination or synchronisation (Richardson et al., 2007; Richardson, Marsh, & Schmidt, 2005; Schmidt & O'Brien, 1997; Varlet, Coey, et al., 2012).

The current study concerns musical synchronisation, a domain that has proven fruitful in the study of control principles in human movement timing (Repp, 2005; Repp & Su, 2013). Here, synchronisation is primarily *intentional* as in the case of instrumental movements—that is, movements that produce sound such as striking a piano key or moving a viola bow—and occasionally unintentional as seen in ancillary movements—movements not causally linked to producing sound, such as body sway (Davidson, 2012; Nusseck & Wanderley, 2009). In most music performance, there is an explicit goal to synchronise sound via instrumental movements, but ancillary movements are a widely observed phenomenon as well. Intentional synchronisation is achieved through a number of cognitive and motor processes (Repp & Su, 2013) such as error correction (Chen et al., 2006; Pressing, 1998a), anticipation (Colley et al., 2017; Mills et al., 2015) and using specific trajectories of movement (Balasubramaniam et al., 2004; Doumas & Wing, 2007; Elliott et al., 2009). On the other hand, ancillary coordination in music performance may be an emergent property related to perception-action coupling such that ancillary movements (the action) change with

the expressive content (the perception) of the music (Demos et al., 2014), or related to larger temporal structures in music such as phrases (MacRitchie et al., 2013).

Much of the research on unintentional synchronisation stems from theories of entrainment, which is the phenomenon for two or more periodic processes to become coupled and thereby unfold in a similar if not identical way (Phillips-Silver, Aktipis, & Bryant, 2010; Strogatz & Stewart, 1993). In human movement, unintentional entrainment is often realised through visuo-motor coupling, meaning an individual can see, even peripherally, the environmental rhythm that entrains his or her movements (Richardson et al., 2007; Varlet, Coey, et al., 2012). The specific conditions under which unintentional visuo-motor entrainment occurs vary, but generally it seems that high amplitude (Varlet, Coey, et al., 2012) and velocity (Varlet, Coey, et al., 2014) of movement in an environmental rhythm yield greater entrainment in observers. Furthermore, the visual tracking of periodic movement increases the stability of both unintentional and intentional entrainment (Schmidt et al., 2007). This is probably due to greater "information pick-up" when eye movements are allowed to follow the stimulus compared to fixating on a stationary location (Varlet, Bucci, et al., 2015). These features of visuo-motor coupling are relevant to coordination in a musical ensemble, as there are peripheral visual signals in the movements of co-performers. Therefore, we expect that unintentional entrainment of ancillary movements between coperformers may occur during music performance (D'Ausilio et al., 2012; MacRitchie et al., 2017; Volpe, D'Ausilio, Badino, Camurri, & Fadiga, 2016).

During intentional musical synchronisation—staying in time with co-performers and "keeping a beat"—relevant visual information improves synchrony even when the task is to synchronise with a sound. For example, in a dyadic dance task where partners moved to a metronome, partners showed lower and less variable relative phase (i.e. greater synchrony) when they looked at each other (Miyata et al., 2017). Similarly, pianists tend to look at each

other prior to difficult musical passages in duets (Kawase, 2014), and move their heads more when auditory feedback is reduced, as if to increase interpersonal visual cues (Goebl & Palmer, 2009). In large ensembles, musicians often monitor a conductor who provides a visual beat for the music. Studies have shown that maximal ensemble synchrony in an orchestra correlates with the maximum vertical velocity of the conductor gesture (Luck & Toiviainen, 2006), and that absolute acceleration of the conductor gesture predicts greater synchrony in a synchronisation tapping task (Luck & Sloboda, 2008). This is true regardless of music experience (Luck & Sloboda, 2009), suggesting that conductor kinematics relate to universal perceptual processes.

In previous work, we found evidence that conductor gestures can improve synchrony by showing that participants with or without music experience tap more accurately with music, and showed greater anticipatory timing—that is, predicting beat intervals during tempo-changing passages—when observing a virtual conductor compared to a visual cue with no acceleration (Colley, Varlet, MacRitchie, & Keller, 2018). Such studies suggest that explicit visual cues—either co-performers when playing without a conductor, or a conductor in a large ensemble—are related to, and may improve intentional synchronisation. In this study, we also found that the conductor caused more structured ancillary movements of the head in participants with no musical training, suggesting that musically-relevant visual cues can influence ancillary movements as well as instrumental movements.

Another distinction in how people achieve musical synchronisation is between sensory modalities; individuals rely both on auditory and visual cues and feedback to optimise their performance. Usually, people synchronise taps more accurately with discrete auditory signals than with discrete visual signals (Elliott et al., 2009; Grahn, 2012; Repp, 2005; Repp & Su, 2013). For continuous signals, the opposite is true: synchronisation tends to be more accurate for continuous visual than for continuous auditory cues (Hove, Fairhurst,

Kotz, & Keller, 2013; Hove, Iversen, Zhang, & Repp, 2012). Multimodal (audio-visual) signals yield better synchronisation compared to unimodal signals (Elliott et al., 2010), especially when the auditory component is discrete and the visual component is continuous (Varlet, Marin, et al., 2012). Thus, input from both modalities is likely beneficial to a performing musician. Indeed, it is common practice to ensure members of a small ensemble can see each other, while large ensembles are usually guided by a conductor. These arrangements provide continuous visual cues in addition to discrete auditory feedback that is inherent in the music.

Despite independent investigations into both intentional audio-motor synchronisation and spontaneous visuo-motor synchronisation, it is likely that in a musical context individuals rely on both auditory and visual information—and likely rely on multiple sources of visual information—to produce a synchronised sound. More specifically, auditory signals provide feedback for error correction (Repp, 2002b), while visual cues such as the conductor provide pre-emptive (i.e. before a beat occurs sonically) information about musical timing (Colley et al., 2018; Wöllner et al., 2012), and ancillary movements of co-performers may provide information about their intentions (Davidson & Malloch, 2009; Keller & Appel, 2010; Ragert et al., 2013). Thus, if both the conductor and a co-performer are available in addition to auditory feedback, individual synchrony should be greater than if only one visual cue is available. Therefore, in the present study, we investigated the effect of multiple sources of visual information on both intentional audio-motor synchronisation, and unintentional visuomotor entrainment between dyads in a controlled musical setting. This was to understand how two sources of visual temporal information (the conductor and a partner) can affect two types of movements (instrumental and ancillary).

Specifically, we asked participants of varying musical experience to drum along to a musical pacing signal with a partner, while observing a virtual conductor. The independent

variable was the visual information available to dyads, with three possibilities: 1) participants were seated side-by-side, facing a monitor so both could see each other and the conductor (condition name: Conductor+Partner/Conductor+Partner); 2) they were lined up so the person in front could see the conductor, and the person in back could only see their partner (Conductor/Partner); or 3) they were side-by-side but with a screen between them, so both could see the conductor but not each other (Conductor/Conductor). (Note that conditions are named so each side of the slash describes what one member of a pair can see.) If both sources of visual information—both the conductor and the partner—factor into successful synchronisation, then dyads should show the lowest asynchrony with the music when they see each other and the conductor. Given the results of our previous study discussed above (Colley et al., 2018), we also expect visual cues to affect anticipatory timing, such that individual temporal prediction will be greater when a participant can see the conductor, than when they are in the back of the line-up arrangement. Additionally, to measure the occurrence of unintentional visuo-motor entrainment, we used a cross-spectral coherence analysis (see Richardson et al., 2005; Schmidt & O'Brien, 1997; Varlet, Bucci, et al., 2015) on the head movements of participants, and expected interpersonal coherence of ancillary movements to be highest when partners could see each other. Lastly, if unintentional ancillary coherence is related to intentional synchronisation, then individuals in pairs with high coherence should show lower drumming asynchrony.

## 2. Methods

2.1. *Participants*. We recruited 40 participants (median age = 27, age range 19-50, 19 female) from the Western Sydney University subject pool and the surrounding area using social media posts. Pairs were created as a convenience sample, such that every two participants who signed up for the experiment formed a pair. We checked for familiarity in dyads by asking: "Have you met this person before? If so, how long ago?" Of the 20 pairs, 19

were unacquainted with each other. The remaining pair had been acquainted for four days prior to their participation, as they had just started their academic semester together. Western Sydney students were given course credit for participation, and non-students were paid \$20 for participation. Three participants were left-handed. Eight participants had more than five years of musical training (mean = 12.62 years, range = 5-20 years), and were currently practicing a musical instrument, which is our definition of a musician. The remaining 32 participants qualified as nonmusicians, having three or fewer years of musical training (mean = .94 years, range = 0-3 years) and ceased playing six years before the experiment on average (range = 4-8 years). Of the 20 dyads, 2 of the dyads were composed of two musicians. The rest were mixed, or entirely nonmusicians.

2.2. *Apparatus*. An Alesis Percpad (tapping pad) was used to collect the drumming data in MIDI format. Participants' movements were recorded with a 12-camera Vicon motion capture system at 100 Hz sampling rate, with reflective markers arranged using a custom model with four markers on the head, one on each shoulder, one on the back of the neck, and one on a single shoulder blade. One partner had the right shoulder blade marked, and the other had the left shoulder blade marked to help distinguish them during data processing.

The motion capture recording and the drum recording were synced by sending a serial trigger signal to Nexus (the motion capture software) at the onset of each trial. The experimental procedure (stimuli presentation, trigger signals, and data collection) was programmed using the OpenFrameworks coding environment for C++ on a 2015 MacBook Pro. Auditory stimuli were sent through stereo speakers, and visual stimuli were presented on a 17" monitor with a 60 Hz refresh rate.

2.3. Auditory stimuli. Three short pieces (Colley et al., 2018) of length 2 m 30 s were created (by author IDC) for the experiment using the musical notation software *Musescore*. The intention was to create stimuli with a constant and unambiguous beat, but with some

melodic and harmonic interest to simulate a musical setting. Thus, the only rhythmic values used were quarter notes in the upper two voices (glockenspiel and xylophone) and eighth notes in the lower voice (harp). These instruments were chosen as they had rapid onsets and were voted as the most pleasing MIDI instruments during pilot testing. There were no rests (i.e., silent beats), meaning every beat as defined in 4/4 meter included an audible note in the music. The pitch range was C2-A5, which is well within typical ensemble ranges. Melodies and harmonies were based on basic practices in Western music theory. The xylophone and glockenspiel played complementary melodies, while the harp accompanied with chords. The length was chosen to reflect a typical short piece of music and to allow for reliable analyses of motion capture data, as discussed later.

To create the tempo changes, the Musescore files were exported as MIDI files, which were converted to ASCII format, then edited in Matlab to change the note on/off times. The music started at 120 bpm or a 500 ms inter-onset interval (IOI). This steady tempo phase continued for eight beats, then a tempo change would occur over eight beats, either slowing or accelerating. The direction of change would then reverse to bring the music back to 120 bpm for eight beats. Thus, the location of tempo changes was regular, but the direction of change, and the magnitude of change were randomly generated. There were six rates of change for the tempo changes: +/-10, +/-16, and +/-22 ms per beat. These rates of change were chosen based on pilot test results. After editing in Matlab, the files were saved as MIDI files, then opened in Garageband to set the instruments for each track, and lastly saved in AIFF format.

*2.4. Visual stimulus.* The visual stimulus was a virtual conductor (Colley et al., 2018) in the form of a red circle with 13 mm diameter against a black background. Again, these colours were chosen based on pilot feedback. For a detailed description of how the virtual conductor was made, see Colley et al., 2018. In brief, we recruited three conductors to come

to the motion capture studio. They listened to, and then conducted the music used in the experiment while we recorded their conducting gestures. We made the virtual conductor by averaging the motion capture position coordinates of the three conductors (similar to Wöllner et al., 2012). The resulting average position coordinates determined the trajectory of the virtual conductor.

2.5. Procedure. Participants arrived and signed consent forms. Once both participants in a dyad arrived, the experimenter attached the motion capture markers while explaining the procedure. Pairs were instructed to begin drumming after four count-in drum beats, and asked to continue drumming until the music stopped. There was no electronically generated auditory feedback, just the sound of the participants' sticks hitting the drums.

There were three arrangements to manipulate the available visual cues. They were: 1) partners sitting adjacent, facing the monitor so they could both see the conductor and their partner (Conductor+Partner/Conductor+Partner); 2) lined up, in which one partner sat behind the other so only the person in front could see the conductor, and the person in the back could only see their partner (Conductor/Partner); and 3) sitting adjacent with a screen between partners (Conductor/Conductor). Participants stayed in one arrangement for two trials to reduce the total amount of time needed to rearrange. When the condition changed, the experimenter would move the drumpads into positions marked on the floor. In the line-up condition partners swapped locations after a trial to balance who was in front.

To ensure that pairs watched the virtual conductor, there were "catch letters" (Varlet, Bucci, et al., 2015). At random points throughout a trial a letter would appear on screen and participants were asked to say it out loud. In the line-up condition, only the person in front would say the letter, as the person in the back was blocked from seeing the screen. Although participants had a fixed reference with which to synchronise (the music), they were instructed to "drum along to the music, while staying synchronised with each other as much as

possible." There was a single one-minute long practice trial, then 18 experimental trials divided into three blocks of six trials.

*2.6. Design*. For most of our dependent measures (maximum coherence, individual asynchrony, and anticipatory timing), we used a one-way repeated measures ANOVA, or Friedman's Rank Test (nonparametric) where appropriate. The three levels were the visual cue conditions, which affected what visual information was available to each member of the dyad (see **Figure 3.1**): Conductor+Partner/Conductor+Partner; Conductor/Partner; Conductor/Conductor. However, we also analysed coherence across eight frequency bins and the three visual cue conditions (plus a permuted pair control). For this, we used an 8 x 4 repeated measures design, described in more detail in the *analysis* section.



**Figure 3.1.** A schematic of the three experimental conditions. The conditions were named according to what each individual can see: Conductor+Partner/Conductor+Partner (A), Conductor/Partner (B) or Conductor/Conductor (C).

#### 3. Data analysis

Prior to any repeated measures test, outliers were removed, where an outlier was any value more than two SDs from the mean (for normal distributions) or median (for non-normal distributions) value. All statistical tests were performed in jamovi (jamovi, 2018), an open-source statistical software. Normality was tested using the Shaprio-Wilk test.

*3.1. Drumming data.* As in Experiment 1, we used an information processing approach to analyse the drumming data. We used two measurements specifically: asynchrony, and a *prediction/tracking index* (P/T index, a measure of anticipatory timing). To measure asynchrony, we calculated the difference between the cumulative inter-drum intervals (IDI) in milliseconds and the cumulative musical pulse intervals in the music, converted each difference score to an absolute value, and averaged the values to produce a mean absolute asynchrony score.

Temporal anticipation was quantified using a *prediction/tracking index* (P/T index) using ARIMAX modelling (ARIMA with an external regressor, X). This is conceptually similar to a previously used method involving cross-correlation (Colley et al., 2017; Colley et al., 2018; Pecenka & Keller, 2011; Repp, 1998). However, the ARIMAX method removes autocorrelation that is inherent in the auditory stimuli used in this experiment, which can overestimate individual prediction abilities. The ARIMAX method works as follows: the cumulative drum series is fit to two ARIMA(1,1,0) models. Each model serves to difference the series (to produce inter-drum intervals), and remove the strong lag-1 autocorrelation found in all participants' data (a by-product of synchronising with music that has gradual, linear tempo changes). The two models differ in their use of external regressors. One model uses the lag-0 IOI series (the normal, unchanged series) to predict the IDI series, while a second model uses the lag-1 of the IOI series (a copy of the series that is shifted back by one observation) to predict the IDI series. The coefficient of the lag-0 model is divided by the

coefficient of the lag-1 model to give a P/T score. If participants are anticipating IOIs, then the lag-0 model should produce a higher coefficient than the lag-1 model. Thus, temporal anticipation is high to the extent that the P/T score is greater than 1.

For both asynchrony and P/T, the IDI and IOI series needed to be the same length. If the series lengths did not match (due to missed, or extra beats in the drum sequence for example), we used the following interpolation procedure: both the IDI and IOI series were cumulatively summed. For each musical beat time (the IOI series), a corresponding drum stroke time was identified by finding the closest drum stroke time, within a window of +/-25% of the given IOI (+/- 81 ms for the smallest IOI, +/- 157 ms for the largest. If there were no drum stroke time values within 25% of an IOI, this was considered a missed beat. This resulted in two series of equal length, but with occasional gaps in the drum stroke series. The drum series for each trial was then fit to a custom ARIMA model to identify serial correlations in the trial. This model was then used in a Kalman filter (Chow, Ferrer, & Nesslroade, 2007; Goodwin & Sin, 1984) to impute missed drum stroke times. The imputed values were then inserted into the gaps of the drum stroke series. Drum stroke series missing more than 20% of the total number of beats were discarded. 4.4% of all trials were discarded.

*3.2. Ancillary movements.* Again, we used analytical approaches from the dynamical systems school for our analysis of ancillary movements, which focused on one marker located on the head (the right forehead marker), based on pilot testing and a previous experiment (Colley et al., 2018). To standardize movement volume of participants, the four markers around the hips were averaged to create a centre point for each trial, which was subtracted from the head movement position values. This helped reduce extraneous, global participant movements such as adjusting their seat away from its original position. To reduce processing time of the motion capture data, the 100 Hz recordings were down-sampled to 50 Hz. Next, the series were filtered with a 10 Hz low-pass Butterworth filter.

To assess the degree of entrainment between partners' ancillary head movements we used cross-spectral coherence analysis. We used a custom R Studio script to run this analysis. Cross-spectral coherence measures the consistency of the phase relation between two time series across a range of possible frequencies. As such, cross-spectral coherence provides an index of entrainment that accounts for multiple frequencies of movement, as well as changing frequencies of movement, such as we would expect in our experiment as the stimuli contain tempo changes. The range of frequencies tested was 0.1 Hz-8 Hz, with a window size of 512 data points, and a 50% window overlap. This frequency range was divided into eight bins (0.1-1 Hz, 1.1-2 Hz...7.1-8 Hz), and the average coherence score from each bin was calculated. These eight averaged coherence scores were taken to represent the level of coordination between partners within the range of movement frequencies in each bin. Coherence scores range from 0 (no coordination) to 1 (perfect coordination).

The descriptive analysis of the averaged coherence bins showed several non-normal distributions across all bins and conditions, and these were not corrected when applied with the appropriate data transformations. Also, we were only interested in comparing across conditions at each frequency bin rather than testing for differences across all conditions and bins. Thus, we tested for differences between each pair of conditions by using bootstrapped 95% confidence intervals (Efron & Tibshirani, 1993; Kirby & Gerlanc, 2013) of mean difference scores. The use of bootstrapping (10,000 iterations) addressed the issue of differently skewed distributions, and the use of difference scores maintained the within-subjects design. Multiple comparisons were addressed by dividing the bootstrap alpha level by the number of pairwise comparisons (three pairwise comparisons in this case) similar to a Bonferroni correction. With this technique, a difference between two conditions is indicated by a confidence interval that does not include zero. The magnitude of the difference between two conditions is indicated by the interval's distance from zero.

Our use of cross-spectral coherence was meant to test for visuo-motor entrainment. However, given the task environment it is possible that pairs would show high coherence measurements simply because they are synchronising with the same music. Therefore we used permuted pseudo-pairs as a control comparison. This involves calculating the coherence between a participant on a given trial and their partner from all trials except the given trial, then averaging the coherence score from all of those comparisons. In other words, we measured coherence in the absence of visual coupling. By comparing this to coherence when partners were actually performing together, we can test for differences in coherence between visually-coupled and non visually-coupled partners.

## 4. Results

*4.1. Synchronisation drumming: Asynchrony.* We first tested the effect of visual cues on individual asynchrony (**Figure 3.2A**). We used data from all individual participants, which made the sample size 38 (two outliers removed) as opposed to 20 pairs. The ANOVA was significant, F(2, 74) = 10.4, p < .001,  $\eta^2 = .03$ , such that the Conductor/Partner visual condition yielded higher asynchronies than both the adjacent no-screen condition (p < .05), and the adjacent with-screen condition (p < .001), Bonferroni corrected.

Given the finding that individual asynchronies were higher in the Conductor/Partner condition—which involved sitting lined up—we tested whether there was an effect of an individual's location within the line-up on their asynchronies (**Figure 3.2B**). To do this we compared each participant's asynchronies when in the front of the line-up, to their asynchronies in the back of the line-up (as a reminder, the line-up arrangement was balanced so that pairs alternated who was in front). The distribution of "back asynchronies" was positively skewed and not corrected with a log10 transformation, so we used a Wilcoxon rank test with the hypothesis that participants would have higher asynchronies when sitting in the back of the line-up. This was supported, W(35) = 215, p < .05, Cohen's d = -.29.



**Figure 3.2.** Panel A: The mean absolute value of asynchronies with the music in milliseconds across the three visual cue conditions. Panel B: Asynchrony with the music while in front of the line-up arrangement (Conductor/Partner condition), compared to the back of the line-up. Error bars represent standard error of the mean for both panels.

4.2. Synchronisation drumming: Anticipatory timing. We also analysed the anticipatory timing of participants using the P/T index (**Figure 3.3A**). The distribution of P/T scores was positively skewed for all three conditions, but was corrected by a log-10 transformation. There was an effect of visual cues, F(2, 72) = 6.21, p < .01,  $\eta^2 = .03$ . Post-hoc comparisons showed that the Conductor/Conductor condition yielded significantly higher P/T scores than the Conductor+Partner/Conductor+Partner condition (p < .01, Bonferroni corrected). The Conductor/Partner condition was intermediate, but not significantly different from the other two conditions.

As with the asynchrony analysis, we examined if there was a difference within the Conductor/Partner condition between the front and back placement (**Figure 3.3B**). We used the log-10 transformed data and a paired-samples *t*-test with the hypothesis that the P/T scores would be higher in the front placement, than in the back. However, the test was not significant, t(36) = 1.09, p = .14, d = .18.

4.3. Asynchrony as a function of anticipatory timing. As an exploratory analysis, we checked for a potential distractor effect, such that the person in back of the line-up, if they are performing badly, might negatively influence the person in front. We did this by testing for asynchrony differences within the Conductor/Partner condition between the person in front of the line-up and the person in the back of the line-up, but within two separate conditions: trials where the person in front was the *higher* predictor, and trials where the person in front was the *lower* predictor. When the person in front was the lower predictor of the pair, their asynchrony was lower than the person in back, W(15) = 32.0, p < .05, Cohen's d = ..49. When the person in front was the higher predictor of the pair, there was no difference in asynchrony between the person in front and the person in the back, W(23) = 143.0, p = .43, Cohen's d = .05.



**Figure 3.3.** Panel A: The mean P/T index across the three visual cue conditions. Panel B: P/T index while in the front of the line-up (Conductor/Partner condition), compared to the back of the line-up. Error bars represent standard error of the mean. Note that non-transformed scores are plotted, although the statistical test for P/T Index used log10 transformed values.

We then compared asynchronies of the person in front when they were the higher predictor, to when they were the lower predictor, but there was no difference, W(15) = 60.0, p = .71, Cohen's d = -.20. We did the same comparison for the person in the back, and again there was no difference, W(15) = 85.0, p = .40, Cohen's d = .31.

*4.4. Synchronisation of ancillary movement.* To assess the synchronisation of ancillary head movements within pairs we first measured the maximum cross-spectral coherence across all frequencies of movement (**Figure 3.4**). There was a significant effect of *visual cue*, F(2, 38) = 8.04, p < .001,  $\eta^2 = .13$ , such that the

Conductor+Partner/Conductor+Partner condition yielded higher maximum coherence than the Conductor/Partner condition (p < .01), the Conductor/Conductor condition (p < .05), and the permuted pairs (p < .001), Bonferonni corrected. However, maximum coherence did not differ between the Conductor/Partner and the Conductor/Conductor conditions. The Conductor/Conductor condition showed significantly higher coherence than the permuted pairs (p < .05). The Conductor/Partner condition however was not significantly different than the permuted pair control. Although maximum coherence values were chosen from the entire range of measured frequencies (0.1 Hz to 8 Hz), all maximum coherence values across all conditions fell within a relatively small range of 1.8 Hz to 2.2 Hz.

We then compared average coherence values from eight different frequency bins, across the three conditions (plus the permuted pair control) using 95% confidence intervals of the difference scores between conditions. The descriptive plot of means is shown in **Figure 3.5** while the confidence interval zero-crossings are shown in **Figure 3.6**. We see that coherence was higher in the Conductor+Partner/Conductor+Partner condition than the Conductor/Partner in the ranges of 1.1-2.0 Hz and 4.1-5.0 Hz. At these frequency bins, the Conductor+Partner/Conductor+Partner condition showed higher coherence than the control condition as well. There were no differences in coherence between

Conductor+Partner/Conductor+Partner and Conductor/Conductor at any frequency bin. The remaining comparisons are best summarised visually in Figure 3.6.



coherence values on a scale from 0 to 1, where 0 represents no coordination at any point in a trial, and 1 represents perfect coordination throughout a whole trial. Data for single pairs are presented in dotted grey lines. The sample means for each condition are represented by the solid black line. Error bars represent standard error of the mean.



**Figure 3.5**. Mean coherence values from each measured frequency bin, across the four conditions. Error bars represent standard error of the mean.

4.5. Relating ancillary and instrumental movements. Lastly we tested whether there was an inverse relation between the coherence of partners' head movements and the asynchrony of each individual's drumming to the music within each condition, across the eight frequency bins (**Figure 3.7**). In other words, we expected pairs with high cross-spectral coherence measures to be relatively low in asynchrony with the music. Given the skewed distribution of the asynchronies, we used Kendall's Tau coefficient (a nonparametric correlation). All alpha levels were Bonferroni corrected. Within the Conductor+Partner/Conductor+Partner condition, there were significant negative correlations in the 3.1-4.0 Hz frequency bin (p < .01). For the Conductor/Partner condition, correlations were significant at 2.1-3.0 Hz (p < .01), and 5.1-6.0 Hz (p < .01). Lastly for the Conductor/Conductor condition, correlations were significant 3.1-4.0 Hz (p < .01), 4.1-5.0

Hz (p < .001), and 6.1-7.0 Hz (p < .01). There were no significant correlations between asynchrony and the coherence of permuted pairs.



**Figure 3.6**. The 95% confidence intervals of difference scores between coherence values of the different visual cue conditions. Each sub-plot represents a different frequency bin, with condition comparisons listed on the *x*-axis.



Mean Coherence (coefficient)

**Figure 3.7**. Scatterplots showing the relation between asynchrony with the music and crossspectral coherence in each of the visual cue conditions, across all measured frequency bins. Correlations with permuted data are shown in grey. Fit lines are linear slopes. Red stars denote a significant Kendall's Tau correlation in the actual data (as opposed to the permuted data). Grey stars (of which there are none) would denote significance in the permuted data. The axis scales for all plots are shown in the upper-right plot. (\*\* = p < .01; \*\*\* = p < .001.)

## Discussion

This experiment investigated the influence of two sources of visual information on synchronisation in a dyadic drumming task. Paired participants drummed along to tempochanging music while observing a virtual conductor that was matched to the timing of the music. We manipulated the following sources of visual information: 1) both partners could see the conductor and each other; 2) one partner observed the conductor, while the other observed their partner, and 3) both partners could only see the conductor. The dependent variables were asynchrony with the music, anticipatory timing, and cross-spectral coherence of ancillary head movements.

Previous studies point to the importance of predicting another person's actions in coordination tasks (Konvalinka et al., 2010; Pecenka & Keller, 2011; Saygin & Stadler, 2012; Schubotz, 2007), and our past work has shown that conductor gestures can enhance temporal prediction and synchrony, even in musically untrained individuals (Colley et al., 2018). The results of the present experiment go further by demonstrating that having a common visual cue that contains information about upcoming time intervals can benefit a group, in this case dyads. Specifically, we found that dyads were more in sync with the music when they could see the conductor, who provided pre-emptive information about musical beat onset times. The conductor also slightly improved anticipatory timing, but only when partners could both see the conductor but not each other.

The beneficial effects of the conductor display on synchrony required full access to the display by the dyad and was not transferable from one individual to the other. Pairs showed lower asynchrony relative to the music when both partners could see the conductor, which was in the Conductor+Partner/Conductor+Partner condition, and the Conductor/Conductor condition. When one partner was cut-off from the conductor (the Conductor/Partner condition), asynchrony with the music was higher than the two other conditions (when averaged across all individuals), suggesting that mutual visual access to the conductor is ideal for synchronising with the accompanying music, regardless of whether partners can see each other. This is further supported by the fact that individuals showed lower asynchrony when they sat in the front of the Conductor/Partner condition (i.e. the front of the line, where they could see the conductor) than when they sat in the back. However, seeing each other (the Conductor+Partner/Conductor+Partner condition) did not reduce the

asynchrony with the music in pairs when compared to the Conductor/Conductor condition, where partners both saw the conductor but not each other.

As for the anticipatory timing as measured by the P/T index, participants were better at anticipating tempo changes when they could see the conductor, but not their partner. This was observed in the Conductor/Conductor condition, which yielded higher P/T scores compared to both other conditions. Relatively low anticipatory timing could arise in the other conditions because seeing the conductor and the partner provided incompatible visual information at times, if, for example, the downward trajectories of the conductor and partner's arm were not well matched in their kinematic features (Hove et al., 2010). In this case, action prediction afforded by the conductor may have been disrupted.

Contrary to what might be expected, P/T scores were not higher when individuals were in the front of the line-up compared to when they were in the back, even though synchronisation with the music was better when an individual was in front. Again, this could be a matter of disruption by incompatible information, but this time due to incompatible auditory cues. If the person in the back was performing relatively poorly, then the sound of their drumming could disrupt the temporal anticipation of the person in front. The possibility of disruption is supported by previous findings that rushed (i.e. earlier than expected) beats in pacing signal synchronisation tasks tend to cause tempo drift in the form of rushing (Repp, 2003a, 2004), and *all-sense-all* networks in which dyads can attend to all available sources of timing information produce greater asynchronies and tempo drift (van de Rijt, 2018). In our paradigm, a distractor effect might not influence mean asynchrony given the relatively long trial time (which would cause these perturbations to average out), but distractions would presumably affect the P/T index by making participants' inter-drum interval series less similar to the stimulus inter-onset interval series. Such a distraction effect from a partner could undo the faciliatory effect of the conductor.

The possibility that a co-performer plus a conductor is not beneficial to musical synchronisation has been considered before. More specifically, observational work and mathematical modelling suggest that a conductor is not beneficial to timing in groups of nine or fewer musicians (Rasch, 1988). It could be that a single person benefits from a visual cue as in our previous work (Colley et al., 2018), but introducing a second person adds potentially unreliable audio and visual information. A previous study of interpersonal synchronisation and anticipatory timing showed that pairs of mixed "predictors" (i.e. one person shows relatively high P/T scores, and the other shows low P/T scores) performed intermediately compared to high predictor pairs (superior performance) and low predictor pairs (worse performance; Pecenka & Keller, 2011). It is possible that a similar type of influence occurred in the present study, such that if a person in the back of the line-up was exhibiting poor prediction, it would affect the person in front.

Our exploratory analysis showed that individuals in a pair did not differ in asynchrony when the person in front was the higher predictor. However, if the person in back was the higher predictor, then the person in front performed better than the person in the back. This means the participants who could exclusively see the conductor benefited from the conductor, but only when they were the lower predictor in the pair. However, across the 40 individual participants, there were more cases (24) where the person in front (who could see the conductor) was on average the *higher* predictor than the person in front being on average the *lower* predictor (16 cases), suggesting that the conductor was a robust visual cue for facilitating prediction.

It can be noted that the effect sizes for our synchronisation drumming results were fairly small, and previous experiments involving dyadic synchronisation have found conflicting results regarding the effect of visual cues. In a musically simple drum/pacing signal task, there was no reliable effect of being able to see a partner (Nowicki, Prinz,

Grosjean, Repp, & Keller, 2013). However, numerous studies of performing musician duos have found that visual cues and body movements can improve interpersonal coordination (Bishop & Goebl, 2015, 2018a, 2018b; Kawase, 2014). Our experimental paradigm is arguably intermediate to the experimental control inherent in a pacing signal synchronisation task, and the ecological validity of actual music performance (though it leans more towards experimental control). Thus, there could be an important distinction to be made between simple dyadic synchronisation where visual cues seem to have no effect on synchrony, and actual musical synchronisation where performance success very much depends on visual coupling. This distinction would be more pronounced in large ensembles, where there is greater distance between any two performers and more variety of timbres, causing aural information to be less reliable while visual information becomes more valuable. Our experiment as a middle ground showed effects, albeit small ones, of visual cues in visually mediated audio-motor synchronisation.

While seeing each other in the Conductor+Partner/Conductor+Partner condition did not improve the synchronisation of instrumental movements in dyads, being able to see each other did increase their maximum ancillary coherence. There was no difference in maximum coherence between the Conductor/Conductor and Conductor/Partner conditions. Thus, in this dyadic context, it seems that mutual visual access increases ancillary coherence. If the visual connection between partners is one-way as in the Conductor/Partner condition, there is no difference in coherence compared to when there is no visual coupling (the permuted pairs), suggesting the importance of mutual coupling, rather than one-way coupling (Miyata et al., 2017). However, observing the same visual cue but not each other (Conductor/Conductor; no direct visual coupling) increased coherence relative to the permuted pairs (where there was no visual coupling). Perhaps observing the same visual while also hearing the partner's drum beats was sufficient to induce some degree of coupling between partners. In other words,

coupling between partners was multi-modal, not strictly visual.

We also compared coherence across the visual cue conditions in eight different frequency bins. We did this because our musical and visual stimuli both contained periodicities at multiple frequencies, which is also the case in naturalistic performance (e.g. Eerola, Jakubowski, Moran, Keller, & Clayton, 2018; Walton, Richardson, Langland-Hassan, & Chemero, 2015). As such, participant movements would presumably show different effects at different frequencies. In the lowest frequency bin (0.1 Hz to 1.0 Hz) and the highest frequency bin (7.1 Hz to 8.0 Hz), participants tended to show lower coherence than the permuted pair controls. In the ranges of 5.1 Hz to 6.0 Hz and 6.1 Hz to 7.0 Hz, coherence was no different than in the permuted pairs. These frequencies are well outside of the range of frequencies in the musical beat structure (1.5 Hz to 3.0 Hz), so this lack of spontaneous coordination at the extremes is not surprising. In the 1.1 Hz to 2.0 Hz bin, the Conductor+Partner/Conductor+Partner condition showed higher coherence than the Conductor/Partner, with the Conductor/Conductor condition marginally higher than Conductor/Partner. This particular bin contains 2 Hz, the average frequency of the music. Thus, it seems that ancillary movements whose frequencies are closely related to the musical tempo become more coherent when dyads are observing the same visual stimulus. There was also a difference in conditions in the 4.1 Hz to 5.0 Hz bin, such that the Conductor+Partner/Conductor+Partner and Conductor/Conductor conditions showed higher coherence than Conductor/Partner, which in turn was no different than the control. The significant effect at 4.1 Hz to 5.0 Hz is likely just a harmonic of the beat-rate, reflecting higher order movements that are integer multiples of the pulse rates in the music. Again, there seems to be an increase in coherence when both members of the dyad are observing the same stimulus, but only at frequencies of movement related to the musical pulse. The middle bins (2.1 Hz-3.0 Hz, and 3.1 Hz-4.0 Hz) showed no differences in coherence among the three

experimental conditions. This is probably because beat rates in this frequency range were only heard during tempo-changing passages. As such, participant movements would have been more variable as they attempted to adapt to the changing speed of the music.

In other words, partners will increase ancillary coherence most strongly when they have each other in view, but observing the same moving visual cue without seeing each other also increases ancillary coherence relative to the permuted pairs, albeit to a lesser degree. This is consistent with general findings that visual rhythms, even in the periphery, facilitate movement coordination, sometimes unintentionally (Clayton, 2007; Richardson et al., 2005; Schmidt et al., 2007; Varlet, Bucci, et al., 2015). It is also consistent with the finding that dyads show lower relative phase and lower standard deviation of relative phase when facing each other, but only on downward movements in a synchronisation dancing task (Miyata et al., 2017).

The importance of downward movements for synchronisation, while not explicitly tested in the present study, is a common finding in synchronisation literature (Hove et al., 2010; Luck & Toiviainen, 2006; Miura, Kudo, Ohtsuki, & Kanehisa, 2011), and relevant to our findings. The virtual conductor displayed rapid downward movements just before a musical beat, and of course the drumming movements of participants were faster in the downward direction than upward (Balasubramaniam et al., 2004; Krause, Pollok, & Schnitzler, 2010). Future studies of movement in musical settings might consider manipulating the direction and shape of movements. For example, a conductor gesture might be more effective for temporal accuracy if confined to a single (vertical) plane, compared to the curved shape that is typically used (acknowledging that these shapes typically relate to musical expression, and therefore serve another purpose). Even though there are horizontal movements involved in conducting gestures, most conductors, including the three conductors used for our visual stimulus, tend to mark beats along the vertical plane (Luck & Toiviainen,

2006).

We hypothesized that ancillary movements in musical synchronisation might reinforce temporal information to improve instrumental synchrony when partners could see each other. However, asynchrony of each individual's drumming with the music did not differ when partners could or could not see each other. Instead, instrumental synchrony suffered only when an individual could not see the conductor; the visual presence of their partner did not seem to affect synchrony. This could be because the conductor cue was more reliable, and more directly linked to the musical timing. Thus, participants may have selectively attended to the conductor when available as a means of facilitating their drumming. The effect of a co-performer alone on synchronisation will be tested in a followup experiment.

We found significant negative correlations between ancillary coherence and drumming asynchrony across the three conditions at several frequency bins. This means that individuals in pairs with relatively high coherence values showed relatively low asynchrony in these cases. Importantly, asynchrony did not correlate with coherence values of permuted pairs, so we are reasonably confident that these correlations are not simply due to common pulse rates in the music across all conditions. The only condition to show this relationship near 2 Hz (the average pulse rate of the music) was Conductor/Partner. This is also the condition that showed the greatest asynchrony, and lowest coherence, suggesting that it was the most difficult condition. This relationship between asynchrony and coherence could be explained by the vestibular hypothesis of beat-keeping and beat induction (Phillips-Silver & Trainor, 2008; Todd & Lee, 2015; Trainor et al., 2009). Partners who moved their heads with greater regularity or in direct relation to the music may have had an easier time drumming in sync with the music, as similar vestibular stimulation (i.e. stimulation at the beat rate) would factor into their sense of time, and then their motor output. Of course, the relationship may

not be causal, and individuals who are better at sensorimotor synchronisation might also entrain more readily to environmental rhythms in general.

We opted not to have a baseline condition with no visual cues (no partner, no conductor) as we were more interested in the comparison between concurrent visual cues. We also wanted to simulate a musical setting in which there are multiple forms of visual cues: the conductor, who provides an explicit temporal cue, as well as a co-performer who provides a peripheral or implicit temporal cue. Thus, we cannot claim that our visual cue conditions necessarily benefit synchronisation beyond simply listening without visual cues, but previous studies have shown an advantage for audio-visual integration in sensorimotor synchronisation (Armstrong & Issartel, 2014; Elliott et al., 2010; Grahn, 2012; Miyata et al., 2017). We do, however, claim that partners show greater musical timing accuracy when they observe a common, temporally relevant visual cue compared to a situation where only one person sees the cue. Also, ancillary coherence tends to be higher when partners see each other, particularly at frequencies of movement related to the musical pulse. Lastly, when partners observed different visual cues, higher ancillary coherence was associated with lower instrumental asynchrony, again at the frequency of the musical pulse. This experiment thus provides further evidence for the advantage of a human-like visual metronome in sensorimotor synchronisation (Colley et al., 2018), and demonstrates the emergence of ancillary entrainment at the beat frequency when partners are visually coupled.

# **EXPERIMENT 3**

A manuscript submitted to Music and Science

Does movement amplitude of a co-performer affect individual performance in musical synchronisation?

Ian D. Colley

Manuel Varlet

Jennifer MacRitchie

Peter E. Keller

### Abstract

Interpersonal coordination in musical ensembles often involves multi-sensory cues, with visual information about body movements supplementing co-performers' sounds. Previous research on the influence of movement amplitude of a visual stimulus on basic sensorimotor synchronisation has shown mixed results. Unintentional visuo-motor synchronisation seems to be influenced by amplitude of a visual stimulus, but intentional visuo-motor synchronisation is not. A case of visually-mediated-but not strictly visuomotor-synchronisation is music performance, which involves both unintentional (spontaneously coordinating ancillary body movements with co-performers) and intentional (producing sound on a beat) forms of synchronisation. We asked whether visual cue amplitude would affect nonmusicians' synchronisation in a musical drumming task designed to be accessible regardless of musical experience. Given the mixed prior results, we entertained two competing hypotheses. H1: higher amplitude visual cues will improve synchronisation. H2: different amplitude visual cues will have no effect on synchronisation. Participants observed a human-derived motion capture avatar with three levels of movement amplitude, or a still image of the avatar (essentially a hearing-only condition), while drumming along to the beat of tempo-changing music. The moving avatars were always timed to match the music. We measured temporal asynchrony (drumming relative to the music), predictive timing, ancillary movement fluctuation, and cross-spectral coherence of ancillary movements between the participant and avatar. The competing hypotheses were tested using conditional equivalence testing. This method involves using a statistical equivalence test in the event that standard hypothesis tests show no differences. Our results showed no statistical differences across visual cues types. We conclude that there is negligible effect of visual stimulus movement amplitude on basic musical synchronisation.

## 1. Introduction

In ensemble music performance, musicians use multi-sensory cues to achieve a synchronised sound. Such cues likely include: auditory feedback to reduce asynchronies and asynchrony variability (Chen et al., 2002); intrapersonal somatic cues such as head movements to reinforce a sense of musical meter (Phillips-Silver & Trainor, 2007, 2008); and visual cues to facilitate anticipation of upcoming temporal patterns in the music (Colley et al., 2018). Assuming co-performers in a musical environment can see each other, *intra*personal somatic cues may also become *inter*personal visual cues, such that one person's rhythmic body movements might be seen by another person. Indeed, mutual visual access among partners in our previous work (a dyadic sensorimotor-synchronisation task with musical sequences) was found to improve the synchrony of partners' ancillary head movements, as well as their synchronisation with the target auditory stimulus (Colley, Varlet, MacRitchie, & Keller, in prep).

Ancillary movements generally play a role in communicating a performer's expressive intentions, with larger movements signalling increased expressive intensity (Davidson & Broughton, 2016). However, the specific influence of the size of co-performer movements on synchronisation abilities has not been tested. Presumably this is partly because of the ecological difficulty of studying live music performance. But also, movement especially music-related movements—can be quite complex and exhibit multiple periodicities such as in casual/non-professional dancing (Burger et al., 2014), and can change with interindividual personality differences (Luck, Saarikallio, Burger, Thompson, & Toiviainen, 2014; Lumsden, Miles, Richardson, Smith, & Macrae, 2012). As such, it can be difficult to identify what qualities of the movement should be controlled or manipulated. Furthermore, any benefit of a co-performer on synchronisation depends to some extent on the skill and reliability of the co-performer (Pecenka & Keller, 2011). We focused on the role of range of

motion—or *movement amplitude*—of a high-performing co-performer on one's ability to synchronise with a musical beat. We selected movement amplitude as a factor of interest with the assumption that larger movements would be more noticeable to observers, and therefore more likely to influence movement timing. In other words, we tested whether larger body movements of a very accurate co-performer could improve the synchronisation accuracy of an observer. Also, as long as the co-performer's movements were always matched to the musical beat (which we controlled for), then larger movements would produce higher velocities. Velocity has been shown to be an important factor in visually mediated synchronisation (Colley et al., 2018; Luck & Sloboda, 2008, 2009; Luck & Toiviainen, 2006; Varlet, Coey, et al., 2014)

Studies on pure visuo-motor synchronisation (no audio component) have shown mixed results regarding the effect of amplitude of a periodic visual stimulus on one's ability to synchronise with the stimulus. Participants were found to synchronise forearm movements with an oscillating circle better with larger amplitudes of circle movement, even when the period duration was kept the same (Varlet, Coey, et al., 2012). Additionally, postural movements showed greater phase entrainment with greater environmental stimulus movements (Dijkstra, Schöner, & Gielen, 1994). In both cases, synchronisation with the visual stimulus was considered *unintentional*, meaning participants were spontaneously synchronising their movements, possibly without awareness. On the other hand, research on *intentional* rhythmic synchronisation suggests there is no effect of stimulus amplitude (de Rugy, Oullier, & Temprado, 2008; Peper & Beek, 1998). Similarly, synchronising finger taps with an image of a finger featuring apparent motion was not affected by the amplitude of the apparent motion (Hove & Keller, 2010). Additionally, synchronisation tapping with a virtual conductor was not influenced by the amplitude of conductor gestures (Wöllner et al., 2012).

Overall, there is some evidence that unintentional visuo-motor coordination is affected by stimulus amplitude, but there is also evidence that stimulus amplitude has negligible effects on intentional visuo-motor coordination. The aim of the current study was to test whether movement amplitude of a visual stimulus affects one's ability to synchronise in a musical situation, where synchronisation among co-performers is not purely visuo-motor, but audio-motor as well. Another interesting aspect of musical synchronisation is that synchrony is not necessarily intentional. Certainly the main objective in most music is to match sounds in time, and as such, audio-motor synchronisation among performing musicians is intentional. However, any apparent visuo-motor synchronisation is likely unintentional, or ancillary.

We tested the influence of stimulus amplitude by having research volunteers drum to the beat of some simple music, while observing a virtual co-performer (avatar), whose movements were manipulated to exhibit various amplitudes of motion, but were always matched to the musical beat. We recorded their drumming in order to measure the asynchrony of their drum strokes, and to quantify their predictive timing, which is the ability to anticipate upcoming beat intervals (Colley et al., 2017; Colley et al., 2018). We also motion capture recorded participants during the drumming task (Colley et al., 2018) to measure the synchrony of their ancillary body movement with the avatar using cross-spectral coherence, as well as to quantify the determinism of their ancillary movements using detrended fluctuation analysis (DFA).

Given the mixed prior research, we had two separate hypotheses regarding the effect of avatar movement amplitude on one's ability to synchronise with a musical beat: 1) Based on work on unintentional coordination, temporal asynchronies relative to a musical pacing signal will be lower when participants observe an avatar with a large movement amplitude, compared to avatars with relatively small, or no movement amplitude. 2) Based on work on

intentional coordination, temporal asynchronies will be lower with a moving avatar compared to a still image, but will not change with different movement amplitudes. Our other measure from the musical drumming task was predictive timing. Based on the finding that temporally relevant, biological motion (compared to temporally relevant non-biological motion) facilitates predictive timing (Colley et al., 2018) we structured our hypothesis in a similar manner to the previous hypothesis: 1) predictive timing will be higher when participants observe an avatar with a large movement amplitude, compared to avatars with relatively small, or no movement amplitude. 2) predictive timing will be higher with a moving avatar compared to a still image, but will not change with different movement amplitudes.

Regarding our motion capture measures (cross-spectral coherence and DFA), we also had two possible hypotheses: 1) coherence (between the participant and avatar), and  $\alpha$ DFA will be higher when participants observe avatars with larger movement amplitudes, compared to relatively small movement amplitudes, or no movement. 2) coherence and DFA will be higher with avatars featuring any movement compared to a still image, but will be the same across movement amplitudes. To test these hypotheses, we used the method of conditional equivalence testing (Campbell & Gustafson, 2018).

## 2. Methods

2.1. Participants. Participants (N = 30, 23 male,  $M_{age} = 19$ ) were recruited through Western Sydney University's research participation program, and given course credit for completing the experiment. Participants were accepted regardless of musical experience, as basic music synchronisation as well as visuo-motor synchronisation are not specialised skills, and we were interested in synchronisation abilities in the general population. However, we assessed musical training with a questionnaire. Three participants had more than five years of musical training, and were currently involved in instrumental music performance. Of the remaining 27 participants, 12 people reported having one academic year or less of music education, and 15 people reported having no formal music education.

*2.2. Design.* The main experimental design was repeated measures, with one-factor, which we will call *visual cue*. The factor *visual cue* refers to magnitude of movements in the visual stimulus, and had four levels: normal movement, movement amplified by 100%, movement amplified by 200%, and no movement (control). As a shorthand, the four conditions will be referred to as Regular, Amp1, Amp2, and Still respectively.

2.3. Auditory stimuli: The music with which participants drummed was made for a previous experiment (Colley et al., 2018) and is described in greater detail in the associated paper. The duration of each piece was two minutes (and therefore the trial duration was also two minutes). It was composed using MIDI instruments with short sound envelopes (150-250 ms) so that notes in the melody would not overlap, thereby avoiding ambiguous beat onsets. There was no change in rhythm in any of the three instrument parts, so that the lines of music created a single target pulse stream. The average IOI was 500 ms, but there were tempo changes throughout the music (IOI range: 332-668 ms). There were three pieces of music. All three were similar in style but featured the tempo changes at different times in the music. It should be noted that the tempo changes were randomly generated for each of the three pieces when the stimuli were made but were not randomly generated at each experimental session. In other words, all participants heard the same music. Further details about the music structure, timing, and composition can be found in our previous study (Colley et al., 2018).

2.4. Visual stimuli: The avatar used in the visual stimuli was made by averaging the motion capture recordings of 10 high-performing participants from a previous experiment, in which they drummed to the same music used here. Thus, the avatar's movements were directly related to the music. In order to be included in the averaged avatar, a participant had to have no missed beats, and an average absolute asynchrony below 30 ms for all three pieces
of music. They also had to be right-handed. With 10 of these participants identified, we reduced the data in their recordings by selecting a subset of motion capture markers (see **Figure 4.1** for a depiction) that gave the impression of a human body. We removed the left arms from the motion capture recordings, as the model participants tended to exhibit task-irrelevant movements with the left hand (e.g. scratching their head, or resetting a loose marker). Further details about the averaging procedure used to create the avatars can be found in our previous paper (Colley et al., 2018).

Once the base avatar was made, we manipulated its movement trajectory to create the other visual cue conditions (see Figure 4.1). The Amp1 condition was made by expanding the range of motion of all markers along all spatial axes (x, y, z) by 100%. In other words, the position coordinates of the base avatar were linearly mapped to fit in between new minimum and maximum values. Thus the timing and relative shape of the avatars stayed the same, but the range of motion increased. The same was done for the Amp2 condition, but the range was increased by 200%. The Still condition (control) was an image of the avatar in its first frame of animation.

*2.5. Apparatus.* An Alesis Percpad (tapping pad) was used to collect the drumming data in MIDI format. Participants' movements were recorded with a 12-camera Vicon motion capture system at 100 Hz sampling rate, with reflective markers arranged using a custom model with four markers on the head, one on each shoulder, one on the back of the neck, one on the dominant hand, and one on the right shoulder blade. The motion capture recording and the drum recording were synced by sending a serial trigger signal to Nexus (the motion capture software) at the onset of each trial. The experimental procedure (stimuli presentation, trigger signals, and data collection) was programmed using the OpenFrameworks coding environment for C++ on a 2015 MacBook Pro. Auditory stimuli were sent through stereo speakers, and visual stimuli were presented on a 17" monitor with a 60 Hz refresh rate.

*2.6. Procedure.* Participants received a Study Information and Consent form by email after signing up for the experiment. They were given a paper copy to sign when they arrived for the experiment. Next, with permission from the participant, the experimenter attached motion capture markers to the following body parts: inner-wrist, outer-wrist, index finger (all on the dominant hand), both shoulders, and the head (using an elastic headband with four evenly spaced markers attached). While attaching the markers, the experimenter explained the task and answered questions.



**Figure 4.1**. A schematic of the three moving visual cues. The *Regular* condition was the averaged motion profile of natural movements. *Amp1* increased the range of motion of the *Regular* condition by 100% along the horizontal and vertical planes. *Amp2* increased the range of motion of the *Regular* condition by 200%. The *Still* control condition maintained the

image of the avatar shown in this figure for the entire trial (without the scales and arrows). Note that depth of movement was represented by the changing diameters of the circles, but there was very little movement along this axis.

Participants were instructed to "drum along to the beat of the music," to "be aware that the speed of the music would sometimes change," and to "always be watching the visuals on the monitor." In an attempt to ensure participants watched the visual cues, we used catchletters, wherein a letter would appear at the centre of the screen at pseudo-random timepoints during a trial. Participants were told to say these letters out loud so the experimenter could verify that they were observing the screen and reporting the correct letters. Letter appearances were timestamped to assess whether they had any influence on drumming asynchrony (see *Data Analysis* section). No specific instructions regarding movement were given. Instead, participants were told to stand however they felt comfortable throughout the trial, so long as their feet and eyes were facing monitor. There were 24 trials of duration 2 m. Participants had one 30 s practice with no visuals, which they could repeat upon request. There was no electronically generated auditory feedback from the drum, just the sound the drumstick hitting its rubber surface After the experiment, participants were given a short musical background questionnaire to assess their musical training (if any) and music-listening habits.

### 3. Data Analysis

*3.1. Drumming analysis.* The drumming analysis used analytical techniques from the information processing school, meaning we focused on discrete time intervals. To check for unusual influence by the catch letters on asynchronies we used the Seasonal Hybrid Extreme Studentized Deviant (SH-ESD) test on the asynchrony time series. SH-ESD detects outliers in seasonal time series data, "seasonal" meaning the time series has periods of fixed length,

as in our tempo-changing music. SH-ESD is similar to Grubbs test for outliers, but is preferred for time series. To check if the catch letters successfully sustained participant attention the experimenter confirmed that the letter said by the participant matched what appeared on the screen throughout the experiment session.

From our drumming recordings we produced two measurements: asynchrony and predictive timing. Asynchrony was calculated as the average of absolute time differences in milliseconds between the sequence of musical beat intervals (or *inter-onset intervals* [IOIs]) and the sequence of participant drum intervals (or inter-tap intervals [ITIs]). To quantify predictive timing we used the prediction/tracking index (Colley et al., 2017; Pecenka & Keller, 2009a). This measure is the ratio of a prediction coefficient over a tracking coefficient. The prediction coefficient represents the strength of the statistical relationship between the ITI and IOI series. The tracking coefficient is the statistical relationship between the ITI series and the lag-1 IOI series. Thus the prediction coefficient is high if participants are anticipating the changing beat intervals and thereby closely matching the intervals, and the tracking coefficient is high if participants are responding to changing beat intervals one beat later, thereby resembling the lagged IOI series. For asynchrony and P/T Index, we used Grubbs' test to identify outliers.

*3.2. Motion capture analysis.* We used analytical approaches from the dynamical systems school to analyse the motion capture recordings. From our motion capture recordings we produced two measures: cross-spectral coherence, and DFA. For both measures, we used the root-sum-square of the raw motion capture data. This produces a directionless signal that incorporates features from all three spatial planes (x, y, z), and we had no specific hypotheses regarding the direction of participant movements. We reduced the motion capture data further by down-sampling to 50 Hz from 100 Hz, and filtering the resulting signal with a 10 Hz low-pass filter.

Cross-spectral coherence measures the consistency of phase relationships among multiple frequencies in a signal. It produces a value between zero (no synchrony) and one (perfect synchrony). In this case, we are measuring the phase relationships among different frequencies of movement between participants, and the avatar. As there was no movement in the control stimulus (a still image), we used a pseudo-pair control. This means that to analyse control trials, we paired the signal of a participant with the signal of the same participant from a different trial (but a trial featuring the same music). The coherence window size was set at 512, and the overlap size at 50%. The range of measured frequencies was .1 Hz to 8 Hz, and the reported coherence scores are the average of all coherence values from within this range.

DFA quantifies the noise colour of a signal. Briefly, signals can exhibit white noise (random values within a narrow range), pink noise (some degree of predictable patterns; some drift), or Brownian noise (highly predictable pattern; heavy drift). Body sway during passive standing tends to exhibit pink noise (Wang & Yang, 2012). If participants entrain to a rhythmic stimulus, we expect DFA to show values above pink noise, as ancillary body movements become more rhythmic and predictable. The output from DFA is  $\alpha$ , which typically ranges from 0.5 (white noise) to 1.5 (Brownian noise) with 1.0 (pink noise) in between. For both coherence and DFA we again used Grubbs' Test to identify outliers.

3.3. Equivalence test. We used conditional equivalence testing (Campbell & Gustafson, 2018) to address our divergent hypotheses. In traditional hypothesis testing, nonsignificant test statistics indicate that one should not reject the null-hypothesis that two means are equal, but this does not speak to the equivalence of the two or more conditions being compared. In other words, one cannot *accept* the null-hypothesis that two or more means are equal. With conditional equivalence testing, one first uses a standard hypothesis test (in our case, ANOVA). If there are null-results in a comparison of two means of interest, and if it is

relevant to the hypothesis, one then uses an equivalence test to determine whether the means are statistically equal, or if their relationship is inconclusive with the given data.

The equivalence test we used was the Two One-Sided Test (TOST) method (Lakens, Scheel, & Isager, 2018). This involves three basic steps. 1) Setting equivalence bounds [-EQ<sub>low</sub>, EQ<sub>high</sub>]; The equivalence bounds form the range of difference scores that are negligible. The bounds are set to include effect sizes that are considered theoretically equal. If this range is not known or there is no theoretical reason to set a particular set of equivalence bounds, then one uses the smallest detectable effect size given the current data distribution and sample size to set the bounds. 2) Testing whether the difference score of interest falls within the equivalence bounds. This is done by running two one-sided *t*-tests (also called one-tailed tests), with  $H_0 I$  that the mean group difference between conditions is greater than EQ<sub>high</sub>, and  $H_02$  that the mean group difference is less than -EQ<sub>low</sub>. Another way to think of this is as a 90% confidence interval of the estimate of interest (difference scores in this case) that is generated by the two *t*-tests. 3) If both *t*-tests (i.e. the 90% confidence interval of difference score estimates) fall within the equivalence bounds as indicated by significant p values then we reject the null hypotheses that the difference score is either greater than the high equivalence bound, or less than the low equivalence bound, and declare equivalence. If one t-test is non-significant, the confidence interval will exceed the equivalence bounds, and we declare inconclusive results. If both one-sided t-tests are nonsignificant, then the original AVNOVA comparison was significant (this is just a conceptual example, an equivalence test would be unnecessary in this case since the ANOVA was significant).

To set our equivalence bounds we used the data-driven *smallest detectable effect size* method, as we had no theoretical reason to identify *a priori* negligible effect sizes for our measures. We considered basing our equivalence bounds for asynchrony on a just noticeable

difference (JND) for asynchronous beats, but studies on this topic have had mixed results (Drake & Botte, 1993; Halpern & Darwin, 1982), and a JND for asynchrony would depend on IOI size (Friberg & Sundberg, 1995; Lerens, Araneda, Renier, & De Volder, 2014), which is not constant in our stimuli. An asynchrony JND would likely also depend on the acoustical features of a sound (London, Nymoen, Thompson, & Danielson, 2017) and of the room. As such, the *smallest detectable effect size* method of setting equivalence bounds seemed appropriate. The equivalence bounds are shown as dotted lines in **Figure 4.5**.

## 4. Results

*4.1. Asynchrony.* We first checked whether participants succeeded in the catch-letter task. All participants correctly named all letters, so we believe the task was effective. We then tested for outliers in participants' asynchrony series due to the catch-letters. The SH-ESD test showed, on average, 2.6 outlying asynchrony scores for each participant. This is far fewer than the number of letters that appeared in a trial, and only 5 of 78 total outliers across all participants occurred within 500 ms after a letter appearing. As such, we have little reason to believe the letters influenced asynchronies.

Prior to the asynchrony ANOVA, we used a log10 transform as the average asynchrony scores were positively skewed in the *Regular* and *Amp2* conditions. No participants were outliers. The ANOVA showed no statistically significant differences, F(3, 87) = 1.25, p = .30,  $\eta^2 = .01$  (see **Figure 4.2**). Therefore, we used a series of equivalence tests to determine if the different condition comparisons were statistically equal, or inconclusive given the current data. This is best summarized visually in **Figure 4.6**, top row, which shows the 90% confidence intervals that correspond to each TOST comparison. Intervals within the equivalence bounds are statistically equal. We see that asynchrony was statistically equal to Still, and Amp2 to Still. While only marginally non-significant, the remaining comparisons are

considered inconclusive, meaning we cannot conclude a statistical difference or equivalence with the current dataset.

4.2. *P/T Index*. The P/T distributions were positively skewed for all conditions so we used a log10 transform on the data. Three participants were removed as outliers after the transform. The ANOVA showed no statistically significant differences, F(3, 78) = 1.90, p = .14,  $\eta^2 = .02$  (see **Figure 4.3**). The equivalence tests (Figure 4.5, second row) showed equivalence for the following comparisons: Regular to Amp1, Amp1 to Amp2, and Amp2 to Still. Again, the remaining comparisons were inconclusive, but only marginally so.



Figure 4.2. The mean asynchronyVisual Conditionscores expressed in milliseconds.Note that the statistical tests used the log10 transformed data. Error bars represent standarderror of the mean.

4.3. DFA. DFA distributions were all normal. No participants were identified as outliers. DFA values were generally slightly above 1.0 (**Figure 4.4**), and within the range observed in our previous work on ancillary motion (Colley et al., 2018). The ANOVA was not significant, F(3, 87) = 1.90, p = .16,  $\eta^2 = .004$  (**Figure 4.4**). The equivalence tests

showed the following statistical equivalences: Regular and Amp2, Regular and Still, Amp1 and Amp2. The remaining comparisons are inconclusive.

*4.4. Coherence*. The distributions for cross-spectral coherence were normal, and there were no outliers. Coherence values were generally between 0.5 and 0.6 (**Figure 4.5**), which is in line with our previous work (Colley et al., 2019). The ANOVA was significant, F(3, 87) = 531, p < .001,  $\eta^2 = .77$ . A Bonferroni-corrected post-hoc test showed that the pseudo-pair control condition showed lower coherence than all other conditions. There were no other statistical differences. The equivalence test reflected this: there were statistical equivalences for all comparisons of Regular, Amp1, and Amp2. But any of those conditions compared to the Still condition showed confidence intervals well above the equivalence bounds.



**Figure 4.3**. The mean P/T Index scores **Visual Condition** expressed as a ratio of leading/lagging ARMA coefficients (see Methods). Note that the statistical tests used the log10 transformed data, but the natural distributions are shown here. Error bars represent standard error of the mean.



**Figure 4.4**. The mean cross-spectral coherence scores between participants and the avatar (or a pseudo-pair).





error bars represent 90% confidence intervals of difference scores.

# 5. Discussion

This experiment investigated the role of movement amplitude of a visual stimulus in facilitating musical synchronisation and influencing ancillary movements. The visual stimulus of which we manipulated the amplitude was a high-performing virtual co-performer (a motion capture avatar). The rationale for this is that a co-performer can be beneficial to a partner if the co-performer is good at the task (Pecenka & Keller, 2011). Additionally, higher

amplitudes of movement that are timed to a fixed musical sequence produce higher velocities (by moving more distance in the same time), which have been shown to improve musical synchronisation (Colley et al., 2018). Given mixed prior results on movement amplitude and visuo-motor synchronisation, we advanced two hypotheses: if overall musical synchrony (i.e. intentional and unintentional movements) is influenced by the amplitude of co-performer movements, then higher amplitudes of stimulus movement will result in lower asynchrony, and higher coherence; alternatively, if musical synchrony is not influenced by the amplitude of a co-performer, then higher amplitudes of stimulus movement will not produce differences in our dependent measures. We also considered the determinism of ancillary movements (DFA), which is not a measure of synchrony but quantifies the extent to which movements are predictable. If stimulus amplitude influences movements, then we would expect larger amplitudes to produce higher DFA values as movements linked to the musical structure would be relatively predictable. If stimulus amplitude does not influence movements, then we would expect no difference in DFA values across amplitude conditions.

Overall, our results suggest that there is no reliable effect of movement amplitude of a visual stimulus on synchronisation accuracy, predictive timing, ancillary movement fluctuations, or the synchrony of ancillary movements between the participant and the avatar. In fact, a number of comparisons between the moving visual stimulus conditions were statistically equivalent, suggesting that our amplitude manipulation produced three effectively identical stimuli (despite physical differences in the visual displays), and so we have greater support for our second set of hypotheses. What is surprising is that the movement conditions were generally no different than the control condition, in which participants observed a still image. The exception to this was the cross-spectral coherence measure, which showed higher coherence between participants' head movements and moving avatars head movements, than between participants' head movements and a copy of their

own movement from another trial (a pseudo-pair). This finding, alongside the apparent success of the catch-letters, suggests that participants were not ignoring the visual display. If they were not observing the visual cues, then their ancillary coherence in the experimental trials would likely resemble the coherence from the pseudo-pair control.

First, we will discuss the drumming dependent variables: asynchrony and P/T Index. It seems that the intentional synchronisation of our participants was not affected by the moving visual cues, even compared to a still image visual cue. This could be due to participants' generally small amount of training in music, which was reflected in the average absolute asynchrony across conditions (about 45 ms). This is consistent with another synchronisation study that tested nonmusicians with similar tempo-changing stimuli (Mills et al., 2015). For example, *motor experts* (people with experience executing deliberate movements in a given domain) tend to be more perceptually sensitive to gross body movements in their domain. For example, basketball players predict shot success better than referees, who typically observe but do not play the game (Aglioti, Cesari, Romani, & Urgesi, 2008). Similarly, violinists predict tone onsets better than musicians of other instruments when observing video of a violinist performing a cueing motion, a movement meant to help observers predict a tone onset (Wöllner & Canal-Bruland, 2010). More recent work has shown that gestures can effectively convey a beat and tempo in musical duos, but only expert musicians were tested, and musicians with more ensemble experience synchronised better (Bishop & Goebl, 2018a). In another study, musicians were generally able to perceive audiovisual asynchronies in musical performance videos, but pianists showed more perceptual sensitivity when observing other pianists (Bishop & Goebl, 2018b). Given the results of these studies, musical expertise may be beneficial for integrating temporal information from a moving body. This is also seen in the fact that musicians in one study only looked at the conductor 28% of the time, for less than one second each time (Fredrickson, 1994),

suggesting they have trained the ability to receive temporal information from brief glances. Only three of our participants had extensive musical training, and only two had ensemble training, meaning the sample was mostly nonmusicians. The three musicians' asynchrony scores were in the lowest four values of the sample, so they were performing relatively well. However, they did not qualify as outliers so we have no reason to treat them as a separate group. Furthermore, removing the three musicians from the sample (resulting in N = 27) did not change the significance of the results of the hypothesis tests. As such, the participants may have observed the stimuli as instructed, but may not have been able to extract relevant temporal information from a full upper-body display, which had multiple moving parts. In other words, participants did not have experience watching a complex rhythmic stimulus to form a temporal prediction.

Expanding on this, a previous study showed that a video of a conductor (from the waist up, similar to our avatars) yielded more precise tapping than a video of a metronome for musicians, but not nonmusicians. In the same study, neural activation in the superior frontal gyrus correlated positively with the amount of time spent practicing with a conductor (Ono, Nakamura, & Maess, 2015). Both groups performed the same in the metronome condition, perhaps because the metronome had a single moving part that corresponds directly to the beat. A previous study of ours (Colley et al., 2018) showed that both musicians and nonmusicians benefited from a virtual conductor, which was presented as a single moving circle. This suggests that visual cues for intentional synchronisation are most effective for the general population if they are kept simple (i.e. one moving part). Complex whole-body movements likely require training to analyse in real time. Indeed, it has been shown that body movements exhibit multiple periodicities when dancing (Burger et al., 2014; Su, 2016), and that tracking multiple moving objects simultaneously complicates action prediction (Atmaca et al., 2013). Thus, segments of the body that move in relation to a musical beat might be

perceived as individually moving parts rather than as a whole phase-locked system, which in turn might depreciate the value of a visual cue. Future studies might test this explicitly by manipulating the number of visible limbs/moving parts in an avatar, and comparing synchronisation performance between ensemble musicians, solo musicians, and nonmusicians.

Our motion capture results reinforced one common finding: individuals tend to entrain their movements to a visual rhythm (Clayton, 2007; Kotz et al., 2014; Schmidt et al., 2007; Schmidt & Turvey, 1994; Varlet, Bucci, et al., 2015). But this seemingly unintentional visuo-motor entrainment of the head does not appear to be increased by the amplitude of the visual rhythm, at least in a multi-sensory context such as music performance. But again, this may be a matter of expertise, such that experienced ensemble musicians would be more likely to show greater ancillary movement coherence with the amplified avatars. As for the fluctuations of movements as measured by DFA, there was no difference across conditions. Importantly, participants' DFA scores for all conditions were centred just above 1.0, suggesting that people tended to move with little more structure than passive standing balance (Blázquez et al., 2009). We expected the amplitude manipulation to increase DFA scores, indicating more rhythmically structured movements of the participants. If our participants were in fact unable to extract temporal information from the avatars, then they may have neglected the visual information entirely as it was deemed unreliable (Elliott et al., 2010).

Finally, it should be noted that our sample came from a healthy population. However, an individual's ability to control periodic movements can be impaired if afflicted with a motor disorder such as Parkinson's Disease (Hove et al., 2012; Nombela et al., 2013). Research on rehabilitation in Parkinson's Disease has shown that external rhythmic cues both auditory and visual—can restore some functionality to patients (Hove & Keller, 2015;

Ghai, Ghai, Schmitz, & Effenberg, 2018). The moving visual stimuli presented in this experiment might provide some benefit to patients with movement disorders where healthy participants received no advantage relative to the control stimulus.

With our results, we conclude that there is no reliable effect of co-performer movement amplitude on intentional or unintentional synchronisation in music, at least in a largely nonmusician sample. We draw this conclusion based not only on statistically nonsignificant differences, but on several statistically equivalent comparisons as well. Future studies of visuo-motor and audio-visuo-motor synchronisation should consider the influence of expertise, especially in musical synchronisation. Other possible variables of interest are the complexity or richness of the musical material (e.g., in potential for expressive variation) and stimulus movement (as measured by the number of moving parts or distinct movement frequencies). Musical expertise and complexity may be influential to the extent that ancillary movements play a greater role in providing cues for flexibly aligning expressive performance parameters than in facilitating strictly synchronised timing (Keller, 2014). But given our current results, it seems that human-derived virtual co-performers provide negligible benefit to basic musical timing abilities of the general population.

#### **GENERAL DISCUSSION**

### 1. A Review of the Results

The three experiments forming this thesis examined three types of visual cues that are common in music performance: a conductor, a co-performer, and the combination of the two. This was done using a newly devised experimental procedure that builds on the classic sensorimotor synchronisation tapping task and analytical methods from the information processing and dynamical systems literature. Overall, results were generally positive in that they confirmed the basic hypothesis that visual cues—particularly a conductor—can reduce temporal asynchronies and improve anticipatory timing in the general population, not just musicians during musical synchronisation. A live co-performer seems to offer a small benefit to the synchronisation of ancillary movements, but a virtual, pre-programmed co-performer offers no apparent benefit. The role of ancillary head movements was less clear, though the common finding that two moving bodies tend to entrain if they are visually coupled was upheld in Experiments 2 and 3. Meanwhile, the use of DFA in Experiment 1 provided evidence that Brownian noise movement structures-that is, deterministic or rigid movements—are associated with relatively poor synchronisation. This was contrary to the expectation that more structured movements would be associated with lower asynchrony, but is in line with previous findings that Brownian-noise structures (rigid movements) tend to be associated with relatively poor performance in a number of tasks.

A principal finding from these experiments is that conductor kinematics are effective for improving synchronisation by facilitating temporal anticipation (Experiments 1 and 2). This was found for both musicians and nonmusicians, and is presumably due to an informative velocity profile, which clearly marks upcoming pulses (Balasubramaniam et al., 2004; Varlet, Coey, et al., 2014). More generally, it suggests that dynamic visual signals help form accurate temporal predictions when an auditory pacing signal is irregular but

predictable. The limit to this visual cue advantage could be tested by producing more complex auditory sequences that are less predictable. But this project has provided initial evidence that a conductor—or more generally, visual rhythms with high acceleration—tap into anticipatory timing mechanisms. The conductor's influence on ancillary movements was apparent but not entirely clear. The conductor stimulus increased the determinism<sup>3</sup> of nonmusicians' head movements (Experiment 1), but had no effect on cross-spectral coherence of partners' head movements (Experiment 2) despite acting as a common visual rhythm for dyads when they were blocked from seeing each other.

Co-performer ancillary movements are not beneficial to intentional synchronisation, at least in the context of the music used in these experiments. Studies of actual music performance (as opposed to a synchronisation drumming task) have repeatedly shown the importance of visual contact among musicians (Bishop & Goebl, 2015, 2018a, 2018b; Kawase, 2013, 2014). While there is no apparent benefit of co-performer ancillary movements for intentional synchronisation in this context, these ancillary cues do seem to be salient enough as visual rhythms to entrain other ancillary movements (Experiments 2 and 3). Increased ancillary coherence seems to be related to lower asynchrony, but only for certain frequencies of movement (Experiment 2). This relationship between head coherence and intentional drumming synchrony is another interesting novel finding, which suggests that intentional and unintentional movements in music may be related at frequencies other than the average pulse rate of the music. Testing for causality will be an important next step in understanding this relationship. Unfortunately, this is difficult as ancillary movements cannot easily be manipulated without introducing another layer of intention (e.g. "move more/less

<sup>&</sup>lt;sup>3</sup> This refers to the DFA procedure, which is usually interpreted in terms of fractality as described in the introduction. However, DFA is used here as a measure of determinism or structure of movements, since fractality is difficult to intuit and interpret as a musical behaviour.

than usual"), and intentional rhythmic movements may be cognitively demanding (Colley et al., 2017; Jacoby et al., 2013).

Another novel finding involves the use of DFA in Experiment 1. This showed that more structured, deterministic movements are not necessarily useful in music. Instead, the relatively high DFA scores in relatively poor performing nonmusicians might indicate rigidity of movements. Rigidity as measured by DFA is associated with difficulty or relatively poor performance in other domains, as discussed in Experiment 1 (Cohen & Sternad, 2009; Nakayama et al., 2010; Wilson et al., 2008). This project has shown that musical synchronisation is among those behaviours where rigid movements might indicate difficulty with the task. Musical synchronisation might therefore be governed by selforganised criticality (Bak et al., 1987), the idea that variability a system—in this case, the human motor system—will naturally find a balance between randomness and determinism in order to achieve flexibility.

While ancillary movements might be difficult to directly manipulate, one could explore the relationship between head movement fluctuations and task performance by recording head movements across varying levels of difficulty in a sensorimotor synchronisation task. This would have implications for the vestibular hypothesis of beat and meter perception (Phillips-Silver & Trainor, 2008; Todd & Lee, 2015), as rigid movements of the head in particular would trigger mechanisms in the inner-ear.

### 2. Limitations and Future Directions

Despite some novel findings, this project is not without its limitations. Finding an appropriate balance between ecological validity and experimental control is always a challenge when studying complex behaviours such as music (D'Ausilio et al., 2015). For the current project, this was a notable challenge from the start since recording and comparing motion capture and synchrony of full ensembles is not necessarily practical. On the other

hand, using non-musical pacing signals and studying a single person at a time is far from musical synchronisation as it is exhibited in most cultures (Clayton, 2007; Davidson, 2002). The basic procedures and auditory stimuli used in these experiments were meant to find a middle ground, in which participants are always interacting with some other agent—whether it is a real partner, or a virtual one made from motion capture recordings—while attempting to synchronise with actual music.

One obvious limitation of this method is a lack of rhythmic and harmonic variety. The music for these experiments was based on Western classical harmonies, and there were arguably no rhythms in the music, only pulses. This second point in particular is worth following up on, as musical timing rarely involves producing simple pulse streams, but rather non-isochronous sequences (rhythms) based on underlying pulse streams (Grahn, 2012). Research on the repetitive reproduction of two- and three-interval rhythms (that is, non-isochronous tone sequences of length 3 or 4 respectively) has shown that people tend to reproduce the rhythms incorrectly over time by adjusting the intervals to fit a particular ratio (Fraisse, 1956; Repp, London, & Keller, 2012). This effect is true for trained musicians as well (Repp, London, & Keller, 2013). Furthermore, when synchronising with non-isochronous sequences, people tend to vary in their synchronisation strategy (Launay, Dean, & Bailes, 2014). A future line of research might examine whether observing the visual cues studied here can help people maintain the original intervals in a rhythm reproduction task, or perhaps even converge on a strategy in a rhythmic synchronisation task.

Moving closer to ecological validity, one might also study visual cues in relation to rhythms that are heavily syncopated (i.e. very few notes occur on the beat) or expressively timed (i.e. the beat is irregular and fluctuates). Studies suggest that people move more to syncopated music (Witek, Clarke, Wallentin, Kringelbach, & Vuust, 2014; Witek et al., 2017) and synchronise best with the beat rather than in anti-phase (a form of syncopation) of

the beat (Keller & Repp, 2004). One might consider whether added visual cues alter the relationship between syncopated music and both ancillary movements and intentional synchronisation. Regarding expressively timed music, people synchronise worse with music that has a highly variable beat compared to a regular beat (Colley et al., 2017), but musicians seem capable of deriving temporal expectations from expressive timing (Repp, 1998, 1999). Given that visual cues were shown here to improve anticipatory timing, visuals presented during synchronisation with expressively timed music might improve performance. Using rhythmically intricate stimuli or expressively timed music as the synchronisation stimulus would push these methods more towards ecological validity and give a better picture of how visuo-motor timing works in music performance.

On the other hand, the universality of music appreciation and engagement (Hodges, 1996)—such as in social dancing (Burger et al., 2014) or clapping at a concert (Néda & Ravasz, 2000)—is a compelling topic that warrants further exploration by including nonmusicians in studies of musical synchronisation when possible. Indeed Experiment 1 yielded interesting results regarding the ancillary movements of nonmusicians, whereas musicians tended not to move much in this particular musical context. Experiments 2 and 3 on the other hand used mostly nonmusician participants. The results of Experiment 3 suggest that the general population does not benefit from observing a highly accurate co-performer, and as discussed this might be due to a lack of domain-specific training for audio-visuo-motor integration (Ono et al., 2015; Wöllner & Canal-Bruland, 2010). Future studies might test this explicitly by comparing nonmusicians to typically solo-performing musicians (who are musically trained but play alone), and ensemble musicians (who presumably have experience observing co-performers). But these experiments did achieve the goal of analysing music-related movements and temporal processing in a general sample.

Considering that the thesis is primarily focused on the notion of musical synchronisation, an ideal extension of these experiments would involve procedures with more than two participants per session in order to test the dynamics of actual musical performance. This has been done with string quartets (Timmers, Endo, Bradbury, & Wing, 2014; Wing, Endo, Bradbury, et al., 2014). But by applying the stimuli used in the present project, one could test nonmusicians for interpersonal serial dependencies in a group synchronisation task. Comparing nonmusician groups to musician or mixed groups could answer questions of the extent to which task-relevant training determines interdependencies in timing (Papiotis, Marchini, Perez-Carrillo, & Maestre, 2014; Torre, Balasubramaniam, & Delignières, 2010; Wing, 2014), and whether ancillary movements might emerge more readily in group, rather than in single or dyadic contexts as has been studied before (Keller & Appel, 2010; MacRitchie et al., 2013). Analysing motion capture and drum sequences in groups would be very laborious, and was beyond the scope of this project, but would offer exciting extensions of the present findings.

## CONCLUSION

There are many more avenues for investigation in the study of human movement and timing in music. This project considered movement and timing in the context of Western classical ensembles, and focused on relevant visual cues as a potential explanation for how ensembles achieve synchrony. The results reinforced the established importance of visual contact in music, but also showed more generally how music synchronisation can be explained by both emergent and cognitive accounts of movement. Furthermore, there are meaningful statistical relations between these accounts, suggesting that the underlying theoretical differences may be reconcilable. But at this time, one can conclude from this project that visual cues improve the psychological processes related to temporal processing as well as influencing associated motor behaviours, and are a valuable aspect of ensemble music.

#### REFERENCES

- Aglioti, S. M., Cesari, P., Romani, M., & Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nature Neuroscience*, 11(9), 1109-1116. doi:10.1038/nn.2182
- Armstrong, A., & Issartel, J. (2014). Sensorimotor synchronization with audio-visual stimuli: limited multisensory integration. *Exp Brain Res*, 232(11), 3453-3463. doi:10.1007/s00221-014-4031-9
- Atmaca, S., Stadler, W., Keitel, A., Ott, D. V., Lepsien, J., & Prinz, W. (2013). Prediction processes during multiple object tracking (MOT): Involvement of dorsal and ventral premotor cortices. *Brain Behav*, 3(6), 683-700. doi:10.1002/brb3.180.
- Bak, P., Tang, C., & Wiesenfield, K. (1987). Self-organized criticality: An explanation of 1/f noise. *Physical Review Letters*, 59, 381-384.
- Balasubramaniam, R., Wing, A. M., & Daffertshofer, A. (2004). Keeping with the beat: movement trajectories contribute to movement timing. *Experimental Brain Research*, 159(1), 129-134. doi:10.1007/s00221-004-2066-z
- Benoit, C. E., Bella, S. D., Farrugia, N., Obrig, H., Mainka, S., & Kotz, S. A. (2014).
  Musically cued gait-raining improves both perceptual and motor timing in Parkinson's disease. *Frontiers in Human Neuroscience*, 8. doi:ARTN 494
  10.3389/fnhum.2014.00494
- Bishop, L., & Goebl, W. (2014). Context-specific effects of musical expertise on audiovisual integration. *Front Psychol*, *5*, 1123. doi:10.3389/fpsyg.2014.01123
- Bishop, L., & Goebl, W. (2015). When they listen and when they watch: Pianists' use of nonverbal audio and visual cues during duet performance. *Musicae Scientiae*, 19(1), 84-110.

- Bishop, L., & Goebl, W. (2018a). Beating time: How ensemble musicians' cueing gestures communicate beat position and tempo. *Psychology of Music*, *46*(1), 84-106.
- Bishop, L., & Goebl, W. (2018b). Communication for coordination: Gesture kinematics and conventionality affect synchronization success in piano duos. *Psychological Research*, 82, 1177-1194.
- Blázquez, M. T., Anguiano, M., Saavedra, F. A., Lallena, A. M., & Carpena, P. (2009). Study of human postural control system during quiet standing using detrended fluctuation analysis. *Physica A*, 388, 1857-1866. doi:10.1016/j.physa.2009.01.001
- 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. doi:ARTN 903

10.3389/fnhum.2014.00903

- Campbell, H., & Gustafson. (2018). Conditional equivalence testing: An alternative remedy for publication bias. *PLoS One, 13*(4). doi:e0195145
- Castellano, G., Mortillaro, M., Camurri, A., Volpe, G., & Scherer, K. (2008). Automated analysis of body movement in emotionally expressive piano performance. *Music Perception, 26*, 103-120.
- Chen, H. Y., Wing, A. M., & Pratt, D. (2006). The synchronisation of lower limb responses with a variable metronome: The effect of biomechanical constraints on timing. *Gait & Posture*, 23(3), 307-314. doi:10.1016/j.gaitpost.2005.04.001
- Chen, Y. Q., Repp, B. H., & Patel, A. D. (2002). Spectral decomposition of variability in synchronization and continuation tapping: Comparisons between auditory and visual pacing and feedback conditions. *Human Movement Science*, 21(4), 515-532. doi:Pii S0167-9457(02)00138-0

Doi 10.1016/S0167-9457(02)00138-0

- Chow, S.-M., Ferrer, E., & Nesslroade, J. R. (2007). An unscented Kalman Filter Approach to the estimation of nonlinear dynamical systems models. *Multivariate Behavioral Research, 42*(2), 283-321.
- Clarke, E., & Cook, N. (2004). *Empirical musicology: Aims, methods, prospects*. New York City, U.S.A.: Oxford University Press.
- Clayton, M. R. L. (2007). Observing entrainment in music performance: Video-based observational analysis of Indian musicians' tanpura playing and beat marking.
   *Musicae Scientiae*, 11(1), 27-59. doi: 10.1177/102986490701100102
- Cohen, R. G., & Sternad, D. (2009). Variability in motor learning: Relocating, channelling and reducing noise. *Experimental Brain Research*, *193*(1), 69-83.
- Colley, I. D., Keller, P. E., & Halpern, A. R. (2017). Working memory and auditory imagery predict sensorimotor synchronization with expressively timed music. *Quarterly Journal of Experimental Psychology*, *71*(8), 1781-1796.
   doi:10.1080/17470218.2017.1366531
- Colley, I. D., Varlet, M., MacRitchie, J., & Keller, P. E. (2018). The influence of visual cues on temporal anticipation and movement synchronization with musical sequences. *Acta Psychologica*.
- Colley, I. D., Varlet, M., MacRitchie, J., & Keller, P. E. (in prep). Dyadic drumming paper. *TBD*.
- D'Ausilio, A., Badino, L., Li, Y., Tokay, S., Craighero, L., Canto, R., . . . Fadiga, L. (2012).
   Leadership in orchestra emerges from the causal relationships of movement kinematics. *PLoS One, 7.* doi:10.1371/journal.pone.0035757
- D'Ausilio, A., Novembre, G., Fadiga, L., & Keller, P. E. (2015). What can music tell us about social interaction? *Trends in Cognitive Sciences*, *19*(3), 111-114.
  doi:10.1016/j.tics.2015.01.005

- Davidson, J. W. (2002). Social and musical coordination between members of a string quartet: an exploratory study. *Psychology of Music, 30*, 186-201.
- Davidson, J. W. (2012). Bodily movement and facial actions in expressive musical performance by solo and duo instrumentalists: Two distinctive case studies.
   *Psychology of Music, 40*(5), 595-633. doi:10.1177/0305735612449896.
- Davidson, J., & Broughton, M.C. (2016). Bodily mediated coordination, collaboration, and communication in music performance. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (pp. 573-595). Oxford, U.K.: Oxford University Press.
- Davidson, J. W., & Malloch, S. (2009). Musical communication: the body movements of performance. In S. Malloch & C. Trevarthen (Eds.), *Communicative Musicality: Exploring the basis of human companionship*. Oxford, U.K.: Oxford University Press.
- de Rugy, A., Oullier, O., & Temprado, J. J. (2008). Stability of rhythmic visuo-motor tracking does not depend on relative velocity. *Experimental Brain Research*, 184, 269-273.
- Delignières, D., & Torre, K. (2011). Event-Based and Emergent Timing: Dichotomy or Continuum? A Reply to Repp and Steinman (2010). *Journal of Motor Behavior*, 43(4), 311-318. doi:10.1080/00222895.2011.588274
- Delignières, D., Torre, K., & Lemoine, L. (2008). Fractal models for event-based and dynamical timers. *Acta Psychologica*, 127(2), 382-397.
  doi:10.1016/j.actpsy.2007.07.007
- Demos, A. P., Chaffin, R., Begosh, K. T., Daniels, J. R., & Marsh, K. L. (2012). Rocking to the beat: effects of music and partner's movements on spontaneous interpersonal coordination. *J Exp Psychol Gen*, 141(1), 49-53. doi:10.1037/a0023843

- Demos, A. P., Chaffin, R., & Kant, V. (2014). Toward a dynamical theory of body movement in musical performance. *Frontiers in Psychology*, *5*, 477. doi:10.3389/fpsyg.2014.00477
- Dijkstra, T. M., Schöner, G., & Gielen, C. C. (1994). Temporal stability of the actionperception cycle for postural control in a moving visual environment. *Experimental Brain Research*, 97, 477-486.
- Doumas, M., & Wing, A. M. (2007). Timing and trajectory in rhythm production. *Journal of Experimental Psychology-Human Perception and Performance*, 33(2), 442-455.
   doi:10.1037/0096-1523.33.2.442
- Drake, C., & Botte, M.-C. (1993). Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. *Perception & Psychophysics*, 54, 277-286.
- Duarte, M., & Zatsiorsky, V. M. (2001). Long-range correlations in human standing. *Physica Letters A*, 283, 124-128.
- Eerola, T., Jakubowski, K., Moran, N., Keller, P.E., & Clayton, M. (2018). Shared periodic performer movements coordinate interactions in duo improvisations. *Royal Society Open Science*, 5(2), 171520. doi: 10.1098/rsos.171520.
- Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. Boca Raton, FL: Chapman & Hall/CRC.
- Elliott, M. T., Welchman, A. E., & Wing, A. M. (2009). Being discrete helps keep to the beat. *Exp Brain Res, 192*(4), 731-737. doi:10.1007/s00221-008-1646-8
- Elliott, M. T., Wing, A. M., & Welchman, A. E. (2010). Multisensory cues improve sensorimotor synchronisation. *European Journal of Neuroscience*, 31(10), 1828-1835. doi:10.1111/j.1460-9568.2010.07205.x

- Elliott, M. T., Wing, A. M., & Welchman, A. E. (2014). Moving in time: Bayesian causal inference explains movement coordination to auditory beats. *Proc Biol Sci,* 281(1786). doi:10.1098/rspb.2014.0751
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*, 429-433. doi:10.1038/415429a
- Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, 8(4), 162-169.
- Fraisse, P. (1956). Les structures rhythmiques [Rhythmic structures]. Louvain, Belgium:Publications Universitaires de Louvain.
- Fredrickson, W. E. (1994). Band musicians' performance and eye contact as influenced by loss of a visual and/or aural stimulus. *Journal of Research in Music Education*, 4, 306-317.
- Friberg, A., & Sundberg, J. (1995). Time discrimination in a monotonic, isochronous sequence. *Journal of the Acoustical Society of America*, *98*(5), 2524-2531.
- Garrido, S., Eerola, T., & K., M. (2017). Group rumination: Social interactions around music in people with depression. *Frontiers in Psychology*, *8*, 490.
- Garrido, S., & Schubert, E. (2015a). Moody melodies: Do they cheer us up? A study of the effect of sad music on mood. *Psychology of Music*, *32*(2), 244-261.
- Garrido, S., & Schubert, E. (2015b). Music and people with tendencies to depression. *Music Perception: An Interdisciplinary Journal, 32*(4), 313-321.
- Ghai, S., Ghai, I., Schmitz, G., & Effenberg, A.O. (2018). Effect of rhythmic auditory cueing on Parkinsonian gait: A systematic review and meta-analysis. *Scientific Reports*, 8(1), 506. Doi:10.1038/s41598-017-16232-5.
- Ginsborg, J., & King, E. (2009). Gestures and glances: The effects of familiarity and expertise on singers' and pianists' bodily movements in ensemble rehearsals. Paper

presented at the European Society for the Cognitive Sciences of Music, Jyväskylä, Finland.

- Goebl, W., & Palmer, C. (2009). Synchronization of timing and motion among performing musicians. *Music Perception*, 26(5), 427-438. doi:10.1525/mp.2009.26.5.427
- Goebl, W., & Palmer, C. (2013). Temporal control and hand movement efficiency in skilled music performance. *PLoS One, 8.* doi:https://doi.org/10.1371/journal.pone.0050901
- Goodwin, G. C., & Sin, K. S. (1984). *Adaptive Filtering Prediction and Control*. Engelwood Cliffs, New Jersey: Prentice-Hall.
- Grahn, J. A. (2012). See what I hear? Beat perception in auditory and visual rhythms. *Experimental Brain Research*, 220(1), 51-61. doi:10.1007/s00221-012-3114-8
- Grahn, J. A., & Schuit, D. (2012). Individual Differences in Rhythmic Ability: Behavioral and Neuroimaging Investigations. *Psychomusicology: Music, Mind, and Brain, 22*, 105-121. doi:10.1037/a0031188.supp
- Halpern, A. R., & Darwin, C. J. (1982). Duration discrimination in a series of rhythmic events. *Perception & Psychophysics*, 31(1), 86-89.
- Hennig, H., Fleischmann, R., Fredebohm, A., Hagmayer, Y., Nagler, J., Witt, A., . . . Geisel,
  T. (2011). The nature and perception of fluctuations in human musical rhythms. *PLoS One, 6.* doi:10.1371/journal.pone.0026457
- Hills, J. M., Ernst, M. O., Banks, M. S., & Landy, M. S. (2002). Combining sensory information: Mandatory fusion within, but not between, senses. *Science*, 298, 1627-1630. doi:10.1126/science.1075396
- Hodges, D. A. (1996). Human musicality. In D. A. Hodges (Ed.), *Handbook of music psychology (2nd ed.)* (pp. 29-68). San Antonia, TX: IMR Press.
- 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. doi:10.1016/j.neuroimage.2012.11.032

- Hove, M. J., Iversen, J. R., Zhang, A., & Repp, B. H. (2012). Synchronization with competing visual and auditory rhythms: bouncing ball meets metronome.
   *Psychological Research*, 77(4), 388-398. doi:10.1007/s00426-012-0441-0
- Hove, M. J., & Keller, P. E. (2010). Spatiotemporal Relations and Movement Trajectories in Visuomotor Synchronization. *Music Perception*, 28(1), 15-26. doi:10.1525/mp.2010.28.1.15
- Hove, M. J., & Keller, P. E. (2015). Impaired movement timing in neurological disorders: rehabilitation and treatment strategies. *Neurosciences and Music V: Cognitive Stimulation and Rehabilitation*, 1337, 111-117. doi:10.1111/nyas.12615
- Hove, M. J., Spivey, M. J., & Krumhansl, C. L. (2010). Compatibility of motion facilitates visuomotor synchronization. *Journal of Experimental Psychology Human Perception* and Performance, 36(6), 1525-1534. doi:10.1037/a0019059
- Hove, M. J., Suzuki, K., Uchitomi, H., Orimo, S., & Miyake, Y. (2012). Interactive rhythmic auditory stimulation reinstates natural 1/*f* timing in gait of Parkinson's patients. *PLoS One*, 7(3), 1-8. doi:10.1371/journal.pone.0032600
- Huys, R., Studenka, B. E., Rheaume, N. L., Zelaznik, H. N., & Jirsa, V. K. (2008). Distinct timing mechanisms produce discrete and continuous movements. *PLoS Comput Biol*, 4(4), e1000061. doi:10.1371/journal.pcbi.1000061
- Inhoff, A. W., & Bisiacchi, P. (1990). Unimanual Tapping during Concurrent Articulation Examining the Role of Cortical Structures in the Execution of Programmed
  Movement Sequences. *Brain and Cognition*, *13*(1), 59-76. doi: 10.1016/0278-2626(90)90040-U

- Iversen, J. R., & Patel, A. D. (2008). The beat alignment test (BAT): Surveying beat processing abilities in the general population. Paper presented at the International Conference on Music Perception and Cognition (ICMPC 10), Sapporo, Japan.
- Jacoby, N., Jakobi, H., Lieder, I., Tishby, N., & Ahissar, M. (2013). Sensorimotor synchronization and working memory. *Journal of Molecular Neuroscience*, 51, S57-S58.
- Jacoby, N., & Repp, B. H. (2012). A general linear framework for the comparison and evaluation of models of sensorimotor synchronization. *Biological Cybernetics*, 106(3), 135-154. doi:10.1007/s00422-012-0482-x
- jamovi, p. (2018). jamovi (Version 0.9) [Computer Software]. Retrieved from https://www.jamovi.org.
- Kantelhardt, J. W., Koscielny-Bunde, E., Rego, H. H. A., Havlin, S., & Bunde, A. (2001).
  Detecting long-range correlations with detrended fluctuation analysis. *Physica A*, 295, 441-454.
- Kawase, S. (2013). Assignment of leadership role changes performers' gaze behavior during piano duo performances. Paper presented at the International Symposium on Performance Science, Brussels, Belgium.
- Kawase, S. (2014). Gazing behavior and coordination during piano duo performance. *Attention, Perception, & Psychophysics, 76*(2), 527-540. doi:10.3758/s13414-013-0568-0.
- Keller, P. E. (2008). Joint Action in Music Performance. *Enacting Intersubjectivity: A Cognitive and Social Perspective on the Study of Interactions, 10*, 205-221.
- Keller, P. E. (2013). Musical ensemble performance: A theoretical framework and empirical findings on interpersonal coordination. Paper presented at the International Symposium on Performance Science, Brussels, Belgium.

Keller, P. E. (2014). Ensemble performance: Interpersonal alignment of musical expression.
In D. Fabian, R. Timmers, & E. Schubert (Eds.), *Expressiveness in music performance: Empirical approaches across styles and cultures* (Vol. 1, pp. 260-282).
Oxford: University Press.

- Keller, P. E., & Appel, M. (2010). Individual differences, auditory imagery, and the coordination of body movements and sounds in musical ensembles. *Music Perception*, 28(1), 27-46. doi:10.1525/mp.2010.28.1.27
- Keller, P. E., & Burnham, D. K. (2005). Musical meter in attention to multipart rhythm. *Music Perception, 22*, 629-661.

Keller, P. E., & Repp, B. H. (2004). When two limbs are weaker than one: Sensorimotor syncopation with alternating hands. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology*, 57(6), 1085-1101. doi:10.1080/02724980343000693.

- Keller, P. E., & Repp, B. H. (2008). Multilevel coordination stability: Integrated goal representations in simultaneous intra-personal and inter-agent coordination. *Acta Psychologica*, 128(2), 378-386. doi:10.1016/j.actpsy.2008.03.012
- King, E. C., & Ginsborg, J. (2011). Gestures and glances: Interactions in ensemble rehearsal.In A. Gritten & K. E. (Eds.), *New perspectives on music and gesture* (pp. 177-201).Aldershot: Ashgate Press.
- Kirby, K. N., & Gerlanc, D. (2013). BootES: An R Package for Bootstrap Confidence Intervals on Effect Sizes. *Behavior Research Methods*, 45(4), 905-1027. doi:10.3758/s13428-013-0330-5.
- Knoblich, G., & Flach, R. (2001). Predicting the effects of actions: Interactions of perception and action. *Psychological Science*, *12*(6), 467-472.

- Komin, N., Murza, A. C., Hernandez-Garcia, E., & Toral, R. (2011). Synchronization and entrainment of coupled circadian oscillators. *Interface Focus*, 1(1), 167-176. doi:10.1098/rsfs.2010.0327
- Konvalinka, I., Vuust, P., Roepstorff, A., & Frith, C. D. (2010). Follow you, follow me:
   Continuous mutual prediction and adaptation in joint tapping. *Quarterly Journal of Experimental Psychology, 63*(11), 2220-2230. doi:Pii 925280376
- Kotz, S. A., Penhune, V., Henry, M., Large, E., Grahn, J. A., & Dalla Bella, S. (2014).
  Entrainment: A domain general cognitive timing mechanism? *International Conference on Timing and Time Perception*, *126*, 24-26. doi:10.1016/j.sbspro.2014.02.304.
- Koul, A., Cavallo, A., Ansuini, C., & Becchio, C. (2016). Doing it your way: How individual movement styles affect action prediction. *PLoS One*, *11*(10), e0165297.
  doi:10.1371/journal.pone.0165297
- Krause, V., Pollok, B., & Schnitzler, A. (2010). Perception in action: The impact of sensory information on sensorimotor synchronization in musicians and non-musicians. *Acta Psychologica*, 133, 28-37.
- Labuta, J. A. (1982). Basic Conducting Techniques. Englewood Cliffs, NJ: Prentice Hall.
- Lacour, M., Bernard-Demanze, L., & Dumitrescu, M. (2008). Posture control, aging, and attention resources: Models and posture-analysis methods. *Clinical Neurophysiology*, 38(6), 411-421.
- Lacquaniti, F., Terzuolo, C., & Viviani, P. (1983). The law relating the kinematic and figural aspects of drawing movements. *Acta Psychologica*, *54*(1), 115-130. doi:https://doi.org/10.1016/0001-6918(83)90027-6

- Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. Advances in Methods and Practices in Psychological Science, 1(2), 259-269.
- Lamoth, C. J. C., Lummel, R. C., & Beek, P. J. (2009). Athletic skill level is reflected in body sway: A test case for accelometry in combination with stochastic dynamics. *Gait & Posture, 29*, 546-551.
- Large, E. W. (2008). Resonating to Musical Rhythm: Theory and Experiment. In S. Grondin (Ed.), *The Psychology of Time*: Emerald Group Publishing Limited.
- Large, E. W., Herrera, J. A., & Velasco, M. J. (2015). Neural networks for beat perception in musical rhythm. *Frontiers in Systems Neuroscience*, 9, 159.
   doi:10.3389/fnsys.2015.00159.
- Launay, J., Dean, R. T., & Bailes, F. (2014). Evidence for multiple strategies in off-beat tapping with anisochronous stimuli. *Psychological Research*, 78(5), 721-735. doi:10.1007/s00426-013-0513-9.
- Lerens, E., Araneda, R., Renier, L., & De Volder, A. G. (2014). Improved beat asynchrony detection in early blind individuals. *Perception, 43*, 1083-1096. doi:10.1068/p7789
- Loehr, J. D., Large, E. W., & Palmer, C. (2011). Temporal coordination and adaptation to rate change in music performance. *Journal of Experimental Psychology Human Perception and Performance*, 37(4), 1292-1309. doi:10.1037/a0023102
- Logan, G. D. (2003). Executive control of thought and action: In search of the wild homunculus. *Current Directions in Psychological Science*, *12*(2), 45-48. doi: 10.1111/1467-8721.01223
- London, J. (2012). *Hearing in time: Psychological aspects of musical meter*. New York: Oxford University Press.

- London, J., Nymoen, K., Thompson, M., D.L., C., & Danielson, A. (2017). Where is the beat in that note? Comparing methods for identifying the p-center of musical sounds.
  Paper presented at the 16th Rhythm Production and Perception Workshop, Birmingham, U.K.
- Luck, G., & Nte, S. (2008). An investigation of conductors' temporal gestures and conductormusician synchronization, and a first experiment. *Psychology of Music, 36*, 81-99.
- Luck, G., Saarikallio, S., Burger, B., Thompson, M., & Toiviainen, P. (2014). Emotiondriven encoding of music preference and personality in dance. *Musicae Scientiae*, 18, 307-323.
- Luck, G., & Sloboda, J. (2007). An investigation of musicians' synchronization with traditional conducting beat patterns. *Music Performance Research, 1*, 26-46.
- Luck, G., & Sloboda, J. (2008). Exploring the Spatio-Temporal Properties of Simple
   Conducting Gestures using a Synchronization Task. *Music Perception: An Interdisciplinary Journal, 25*(3), 225-239. doi:10.1525/mp.2008.25.3.225
- Luck, G., & Sloboda, J. (2009). Spatio-Temporal Cues for Visually Mediated Synchronization. *Music Perception*, 26(5), 465-473. doi:10.1525/mp.2009.26.5.465
- Luck, G., & Toiviainen, P. (2006). Ensemble musicians' synchronization with conductors' gestures: An automated feature-extraction analysis. *Music Perception, 24*, 189-200.
- Lumsden, J., Miles, L. K., Richardson, M. J., Smith, C. A., & Macrae, C. N. (2012). Who syncs? Social motives and interpersonal coordination. *Journal of Experimental Social Psychology*, 48(3), 746-751. doi:10.1016/j.jesp.2011.12.007
- MacRitchie, J., Buck, B., & Bailey, N. J. (2013). Inferring musical structure through bodily gestures. *Musicae Scientiae*, *17*, 86-108. doi:10.1177/1029864912467632
- MacRitchie, J., Varlet, M., & Keller, P. E. (2017). Embodied expression through entrainment and co-representation in musical ensemble performance. In M. Lesaffre, P. J. Maes, &
M. Leman (Eds.), *The Routledge companion to embodied music interaction* (pp. 150-159). New York: Routledge.

- Madison, G., & Delignéres, D. (2008). Auditory feedback affects the long-range correlation of isochronous serial interval production: Support for a closed-loop or memory model of timing. *Experimental Brain Research*, 193(4), 519-527.
- Madison, G., Karampela, O., Ullén, F., & Holm, L. (2013). Effects of practice on variability in an isochronous serial interval production task: asymptotical levels of tapping variability after training are similar to those of musicians. *Acta Psychologica*, 143, 119-128.
- Maes, P. J., Wanderley, M. M., & Palmer, C. (2015). The role of working memory in the temporal control of discrete and continuous movements. *Experimental Brain Research*, 233(1), 263-273. doi:10.1007/s00221-014-4108-5
- Manning, F. C., Harris, J., & Schutz, M. (2017). Temporal prediction abilities are mediated by motor effector and rhythmic expertise. *Experimental Brain Research*, 235, 861-871.
- Maruta, J., Heaton, K. J., Kryskow, E. M., Maule, A. L., & Ghajar, J. (2013). Dynamic visuomotor synchronization: Quantification of predictive timing. *Behavior Research Methods*, 45(1), 289-300. doi:10.3758/s13428-012-0248-3
- Merchant, H., Grahn, J., Trainor, L., Rohrmeier, M., & Fitch, W. T. (2015). Finding the beat:
  a neural perspective across humans and non-human primates. *Philosophical Transactions of the Royal Society B-Biological Sciences, 370*(1664), 91-106.
  doi:UNSP 2014009310.1098/rstb.2014.0093

Michon, J. A. (1967). Timing in Temporal Tracking. Leiden University, The Netherlands.

Mills, P. F., van der Steen, M. C., Schultz, B. G., & Keller, P. E. (2015). Individual differences in temporal anticipation and adaptation during sensorimotor

synchronization. *Timing & Time Perception, 3*(1-2), 13-31. doi:10.1163/22134468-03002040

- Miura, A., Kudo, K., Ohtsuki, T., & Kanehisa, H. (2011). Coordination modes in sensorimotor synchronization of whole-body movement: A study of street dancers and non-dancers. *Human Movement Science*, 30, 1260-1271.
- Miyata, K., Varlet, M., Miura, A., Kudo, K., & Keller, P. E. (2017). Modulation of individual auditory-motor coordination dynamics through interpersonal visual coupling. *Scientific Reports*, 7. doi:10.1038/s41598-017-16151-5
- Nakayama, Y., Kudo, K., & Ohtsuki, T. (2010). Variability and fluctuation in running gait cycle of trained runners and non-runners. *Gait & Posture*, *31*, 331-335.
- Néda, Z., & Ravasz, E., Brechet, Y., Vicsek, T., Barabási, A.-L. (2000). The sound of many hands clapping. *Nature*, 403, 849-850.
- Nessler, J. A., & Gilliland, S. J. (2009). Interpersonal synchronization during side by side treadmill walking is influenced by leg length differential and altered sensory feedback. *Human Movement Science*, *28*(6), 772-785.
- Nombela, C., Hughes, L. E., Owen, A. M., & Grahn, J. A. (2013). Into the groove: Can rhythm influence Parkinson's disease? *Neuroscience & Biobehavioral Reviews*, *37*, 2564-2570.
- Nowicki, L., Prinz, W., Grosjean, M., Repp, B. H., & Keller, P. (2013). Mutual adaptive timing in interpersonal action coordination. *Psychomusicology: Music, Mind, and Brain, 23*, 6-20.
- Nusseck, M., & Wanderley, M. M. (2009). Music and motion—How music-related ancillary body movements contribute to the experience of music. *Music Perception, 26*, 335-353.

- Ono, K., Nakamura, A., & Maess, B. (2015). Keeping an eye on the conductor: neural correlates of visuo-motor synchronization and musical experience. *Front Hum Neurosci, 9*, 154. doi:10.3389/fnhum.2015.00154
- Oullier, O., & Kelso, J. A. S. (2009). Social coordination from the perspective of coordination dynamics. In R. A. Meyers (Ed.), *Encyclopedia of complexity and systems sciences* (pp. 8198-8212). Berlin, Germany: Springer-Verlag.
- Papiotis, P., Marchini, M., Perez-Carrillo, A., & Maestre, E. (2014). Measuring ensemble interdependence in a string quartet through analysis of multidimensional performance data. *Frontiers in Psychology*. doi:10.3389/fpsyg.2014.00963.
- Pecenka, N., & Keller, P. E. (2009a). Auditory pitch imagery and its relationship to musical synchronization. *Annals of the New York Academy of Sciences*, 1169, 282-286. doi:10.1111/j.1749-6632.2009.04785.x
- Pecenka, N., & Keller, P. E. (2009b). *The relationship between auditory imagery and musical synchronization abilities in musicians*. Paper presented at the 7th Triennial
  Conference of European Society for the Cognitive Sciences of Music (ESCOM 2009),
  Jyväskylä: University of Jyväskylä, Finland.
- Pecenka, N., & Keller, P. E. (2011). The role of temporal prediction abilities in interpersonal sensorimotor synchronization. *Experimental Brain Research*, 211(3-4), 505-515. doi:10.1007/s00221-011-2616-0
- Peng, C.-K., Buldyrev, V., Havlin, S., Simons, M., Stanley, H. E., & Goldberger, A. L.
  (1994). Mosaic organization of DNA nucleotides. *Physical Review E, 49*, 1685-1689.
  doi:10.1103/PhysRevE.49.1685
- Peng, C.-K., Havlin, S., Stanley, H. E., & Goldberger, A. L. (1995). Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos, 5*, 82-87. doi:10.1063/1.166141

- Penhune, V. B., & Zatorre, R. J. (1998). Cerebellar contributions to motor timing: A PET study of auditory and visual rhythm reproduction. *Journal of Cognitive Neuroscience*, 10, 752-765.
- Peper, C. E., & Beek, P. J. (1998). Are frequency-induced transitions in rhythmic coordination mediated by a drop in amplitude. *Biological Cybernetics*, 79, 291-300.
- Phillips-Silver, J., Aktipis, C. A., & Bryant, G. A. (2010). The ecology of entrainment: Foundations of coordinated rhythmic movement. *Music Percept*, 28(1), 3-14. doi:10.1525/mp.2010.28.1.3
- Phillips-Silver, J., Toiviainen, P., Gosselin, N., Piche, O., Nozaradan, S., Palmer, C., & Peretz, I. (2011). Born to dance but beat deaf: A new form of congenital amusia. *Neuropsychologia*, 49(5), 961-969. doi:10.1016/j.neuropsychologia.2011.02.002
- Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science*, *308*, 1430.
- Phillips-Silver, J., & Trainor, L. J. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition*, 105, 533-546.
- Phillips-Silver, J., & Trainor, L. J. (2008). Vestibular influence on auditory metrical interpretation. *Brain and Cognition*, 67(1), 94-102. doi:10.1016/j.bandc.2007.11.007
- Praamstra, P., Turgeon, A., Hesse, C. W., Wing, A. M., & Perryer, L. (2003).
   Neurophysiological correlates of error correction in sensorimotor-synchronization.
   *Neuroimage*, 20(2), 1283-1297. doi:10.1016/S1053-8119(03)00351-3
- Pressing, J. (1998a). Error correction processes in temporal pattern production. *Journal of Mathematical Psychology, 42*, 63-101.
- Pressing, J. (1998b). Referential behavior theory: A framework for multiple perspectives on motor control. *Motor Behavior and Human Skill*, 357-384.

- Pressing, J. (1999a). The referential dynamics of cognition and action. *Psychological Review*, *106*(4), 714-747. doi: 10.1037/0033-295x.106.4.714
- Pressing, J. (1999b). Sources for 1/f noise effects in human cognition and performance. Paideusis - Journal for Interdisciplinary and Cross-Cultural Studies, 2.
- Pressing, J., & Jolley-Rogers, G. (1997). Spectral properties of human cognition and skill. *Biological Cybernetics*, 76(5), 339-347. doi: 10.1007/s004220050347
- Ragert, M., Schroeder, T., & Keller, P. E. (2013). Knowing too little or too much: the effects of familiarity with a co-performer's part on interpersonal coordination in musical ensembles. *Front Psychol, 4*, 368. doi:10.3389/fpsyg.2013.00368
- Rankin, S. K., Large, E. W., & Fink, P. W. (2009). Fractal Tempo Fluctuation and Pulse Prediction. *Music Percept*, 26(5), 401-413. doi:10.1525/mp.2009.26.5.401
- Rasch, R. A. (1979). Synchronization in performed ensemble music. *Acustica*, 43(2), 121-131.
- Rasch, R. A. (1988). Timing and synchronization in ensemble performance. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition* (pp. 70-90). Oxford, UK: Clarendon Press.
- Repp, B. H. (1998). A microcosm of musical expression. I. Quantitative analysis of pianists' timing in the initial measures of Chopin's Etude in E major. *Journal of the Acoustical Society of America, 104*(2), 1085-1100. doi: 10.1121/1.423325.
- Repp, B. H. (1999). Relationships between performance timing, perception of timing perturbations, and perceptual-motor synchronisation in two Chopin preludes.
   *Australian Journal of Psychology, 51*(3), 188-203. doi: 10.1080/00049539908255356.
- Repp, B. H. (2001). Phase correction, phase resetting, and phase shifts after subliminal timing perturbations in sensorimotor synchronization. *Journal of Experimental Psychology: Human Perception and Performance*, 27(3), 600-621.

- Repp, B. H. (2002a). The embodiment of musical structure: effects of musical context on sensorimotor synchronization with complex timing patterns. *Common Mechanisms in Perception and Action*, 19, 245-265.
- Repp, B. H. (2002b). Phase correction in sensorimotor synchronization: Nonlinearities in voluntary and involuntary responses to perturbations. *Human Movement Science*, 21(1), 1-37. doi:Pii S0167-9457(02)00076-3.
- Repp, B. H. (2003a). Phase attraction in sensorimotor synchronization with auditory sequences: Effects of single and periodic distractors on synchronization accuracy. *Journal of Experimental Psychology*, 29, 290-309.
- Repp, B. H. (2003b). Rate limits in sensorimotor synchronization with auditory and visual sequences: The synchronization threshold and the benefits and costs of interval subdivision. *Journal of Motor Behavior*, 35(4), 355-370.
- Repp, B. H. (2004). On the nature of phase attraction in sensorimotor synchronization with interleaved auditory sequences. *Human Movement Science*, *23*, 389-413.
- Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review, 12*(6), 969-992. doi: 10.3758/Bf03206433
- Repp, B. H. (2008). Multiple temporal references in sensorimotor synchronization with metrical auditory sequences. *Psychological Research-Psychologische Forschung*, 72(1), 79-98. doi:10.1007/s00426-006-0067-1
- Repp, B. H. (2010). Sensorimotor synchronization and perception of timing: Effects of music training and task experience. *Human Movement Science*, 29(2), 200-213. doi:10.1016/j.humov.2009.08.002
- Repp, B. H., & Keller, P. E. (2004). Adaptation to tempo changes in sensorimotor synchronization: Effects of intention, attention, and awareness. *The Quarterly Journal* of Experimental Psychology, 57A, 499-521.

- Repp, B. H., Keller, P. E., & Jacoby, N. (2012). Quantifying phase correction in sensorimotor synchronization: Empirical comparison of three paradigms. *Acta Psychologica*, 139(2), 281-290. doi:10.1016/j.actpsy.2011.11.002
- Repp, B. H., London, J., & Keller, P. E. (2012). Distortions in reproduction of two-interval rhythms: When the "attractor ratio" is not exactly 1:2. *Music Perception*, 30(2), 205-223.
- Repp, B. H., London, J., & Keller, P. E. (2013). Systematic distortions in musicians' reproduction of cyclic three-interval rhythms. *Music Perception: An Interdisciplinary Journal*, 30(3), 291-305.
- Repp, B. H., & Moseley, G. P. (2012). Anticipatory phase correction in sensorimotor synchronization. *Human Movement Science*, 31(5), 1118-1136.
  doi:10.1016/j.humov.2011.11.001
- Repp, B. H., & Steinman, S. R. (2010). Simultaneous Event-Based and Emergent Timing:
   Synchronization, Continuation, and Phase Correction. *Journal of Motor Behavior*, 42(2), 111-126.
- Repp, B. H., & Su, Y. H. (2013). Sensorimotor synchronization: a review of recent research (2006-2012). *Psychonomic Bulletin & Review*, 20(3), 403-452. doi:10.3758/s13423-012-0371-2
- Richardson, M. J., Marsh, K. L., Isenhower, R. W., Goodman, J. R., & Schmidt, R. C.
  (2007). Rocking together: dynamics of intentional and unintentional interpersonal coordination. *Human Movement Science*, *26*(6), 867-891.
  doi:10.1016/j.humov.2007.07.002
- Richardson, M. J., Marsh, K. L., & Schmidt, R. C. (2005). Effects of visual and verbal interaction on unintentional interpersonal coordination. *Journal of Experimental Psychology Human Perception and Performance*, 31(1), 62-79.

- Riley, M. A., Bonnette, S., Kuznetsov, N., Wallot, S., & Gao, J. (2012). A tutorial introduction to adaptive fractal analysis. *Frontiers in Physiology*, 3. doi:10.3389/fphys.2012.00371
- Ruiz, M. H., Hong, S. B., Hennig, H., Altenmuller, E., & Kuhn, A. A. (2014). Long-range correlation properties in timing of skilled piano performance: the influence of auditory feedback and deep brain stimulation. *Front Psychol*, *5*, 1030. doi:10.3389/fpsyg.2014.01030
- Saarikallio, S. H. (2008). Music in mood regulation: Initial scale development. *Musicae Scientiae*, *12*(2), 291-309.
- Saygin, A. P., & Stadler, W. (2012). The role of appearance and motion in action prediction. *Psychological Research*, *76*, 388-394. doi:10.1007/s00426-012-0426-z.
- Schmidt, R. C., & O'Brien, B. (1997). Evaluating the dynamics of unintended interpersonal coordination. *Ecological Psychology*, 9(3), 189-206.
- Schmidt, R. C., Richardson, M. J., Arsenault, C., & Galantucci, B. (2007). Visual tracking and entrainment to an environmental rhythm. *Journal of Experimental Psychology Human Perception and Performance*, 33(4), 860-870. doi:10.1037/0096-1523.33.4.860
- Schmidt, R. C., & Turvey, M. T. (1994). Phase-Entrainment Dynamics of Visually Coupled Rhythmic Movements. *Biological Cybernetics*, 70(4), 369-376. doi: 10.1007/Bf00200334
- Schubotz, R. I. (2007). Prediction of external events with our motor system: Towards a new framework. *Trends in Cognitive Sciences*, *11*, 211-218.
- Shaffer, H. (1984). Timing in Musical Performance. *Annals of the New York Academy of Sciences, 423*(May), 420-428. doi: 10.1111/j.1749-6632.1984.tb23450.x

- Shockley, K., Baker, A. A., Richardson, M. J., & Fowler, C. A. (2007). Articulatory constraints on interpersonal postural coordination. *Journal of Experimental Psychology Human Perception and Performance*, 33(1), 201-208.
- Sowinski, J., & Dalla Bella, S. (2013). Poor synchronization to the beat may result from deficient auditory-motor mapping. *Neuropsychologia*, 51(10), 1952-1963.
  doi:10.1016/j.neuropsychologia.2013.06.027
- Stadnitski, T. (2012). Measuring fractality. *Frontiers in Psychology*, *3*, 1-13. doi:10.3389/fphys.2012.00127
- Stergiou, N., & Decker, L. M. (2011). Human movement variability, nonlinear dynamics, and pathology: Is there a connection? . *Human Movement Science*, *30*(5), 869-888.
- Stevens, L. T. (1886). On the time sense. Mind, 11, 393-404.
- Stoffregen, T. A., Fu-Chen, C., Varlet, M., Alcantara, C., & Benoît, G. B. (2013). Getting your sea legs. *PLoS One*, 8, 1-16. doi:10.1371/journal.pone.0066949
- Strogatz, S. H. (1995). Nonlinear Dynamics Ordering Chaos with Disorder. *Nature*, 378(6556), 444-444. doi: 10.1038/378444a0
- Strogatz, S. H., & Stewart, I. (1993). Coupled Oscillators and Biological Synchronization. Scientific American, 269(6), 102-109.
- Thompson, M. R., & Luck, G. (2011). Exploring relationships between pianists' body movements, their expressive intentions, and structural elements of the music. *Musicae Scientiae*, 16(1), 19-40. doi:10.1177/1029864911423457
- Timmers, R., Endo, S., Bradbury, A., & Wing, A. M. (2014). Synchronization and leadership in string quartet performance: a case study of auditory and visual cues. *Frontiers in Psychology, 5.* doi:ARTN 64510.3389/fpsyg.2014.00645

- Todd, N. P., & Lee, C. S. (2015). The sensory-motor theory of rhythm and beat induction 20 years on: a new synthesis and future perspectives. *Frontiers in Human Neuroscience*, 9, 444. doi:10.3389/fnhum.2015.00444
- Toiviainen, P., Luck, G., & Thompson, M. R. (2010). Embodied meter: Hierarchical eigenmodes in music-induced movement. *Music Perception, 28*(1), 59-70. doi:10.1525/Mp.2010.28.1.59
- Torre, K., & Balasubramaniam, R. (2009). Two different processes for sensorimotor synchronization in continuous and discontinuous rhythmic movements. *Experimental Brain Research*, 199(2), 157-166. doi:10.1007/s00221-009-1991-2.
- Torre, K., Balasubramaniam, R., & Delignières, D. (2010). Oscillating in Synchrony with a Metronome: Serial Dependence, Limit Cycle Dynamics, and Modeling. *Motor Control, 14*(3), 323-343.
- Torre, K., Balasubramaniam, R., Rheaume, N., Lemoine, L., & Zelaznik, H. N. (2011). Longrange correlation properties in motor timing are individual and task specific. *Psychonomic Bulletin & Review*, 18(2), 339-346. doi:10.3758/s13423-011-0049-1
- Torre, K., & Delignières, D. (2008). Distinct ways of timing movements in bimanual coordination tasks: Contribution of serial correlation analysis and implications for modeling. *Acta Psychologica*, 129(2), 284-296. doi:10.1016/j.actpsy.2008.08.003
- Trainor, L. J., Gao, X., Lei, J., Lehtovaara, K., & Harris, L. R. (2009). The primal role of the vestibular system in determining musical rhythm. *Cortex*, *45*, 35-43.
- Treffner, P. J., & Turvey, M. T. (1993). Resonance Constraints on Rhythmic Movement. Journal of Experimental Psychology-Human Perception and Performance, 19(6), 1221-1237. doi: 10.1037//0096-1523.19.6.1221
- van de Rijt, A. (2018). All-sense-all networks are suboptimal for sensorimotor synchronization. *PLoS One, 13*(8). doi:https://doi.org/10.1371/journal.pone.0202056

- van der Steen, M. C., Jacoby, N., Fairhurst, M. T., & Keller, P. E. (2015). Sensorimotor synchronization with tempo-changing auditory sequences: Modeling temporal adaptation and anticipation. *Brain Research*, *1626*, 66-87.
- van der Steen, M. C., & Keller, P. E. (2013). The ADaptation and Anticipation Model
  (ADAM) of sensorimotor synchronization. *Frontiers in Human Neuroscience*, 7.
  doi:ARTN 25310.3389/fnhum.2013.00253.
- Varlet, M., Bardy, B. G., Chen, F. C., Alcantara, C., & Stoffregen, T. A. (2015). Coupling of postural activity with motion of a ship at sea. *Experimental Brain Research*, 233(5), 1607-1616.
- Varlet, M., Bucci, C., Richardson, M. J., & Schmidt, R. C. (2015). Informational constraints on spontaneous visuomotor entrainment. *Human Movement Science*, 41, 265-281. doi:10.1016/j.humov.2015.03.011
- Varlet, M., Coey, C. A., Schmidt, R. C., Marin, L., Bardy, B. G., & Richardson, M. J. (2014).
  Influence of stimulus velocity profile on rhythmic visuomotor coordination. *Journal* of Experimental Psychology Human Perception and Performance, 40(5), 1849-1860. doi:10.1037/a0037417
- Varlet, M., Coey, C. A., Schmidt, R. C., & Richardson, M. J. (2012). Influence of stimulus amplitude on unintended visuomotor entrainment. *Human Movement Science*, 31(3), 541-552. doi:10.1016/j.humov.2011.08.002
- Varlet, M., Marin, L., Issartel, J., Schmidt, R. C., & Bardy, B. G. (2012). Continuity of visual and auditory rhythms influences sensorimotor coordination. *PLoS One*, 7(9), e44082. doi:10.1371/journal.pone.0044082
- Varlet, M., & Richardson, M. J. (2015). What would be Usain Bolt's 100-meter sprint world record without Tyson Gay? Unintentional interpersonal synchronization between the

two sprinters. Journal of Experimental Psychology Human Perception and Performance, 41(1), 36-41.

- Varlet, M., Stoffregen, T. A., Chen, F. C., Alcantara, C., Marin, L., & Bardy, B. G. (2014).
  Just the sight of you: Postural effects of interpersonal visual contact at sea. *Journal of Experimental Psychology Human Perception and Performance*, 40(6), 2310-2318.
- Vesper, C., Butterfill, S., Knoblich, G., & Sebanz, N. (2010). A minimal architecture for joint action. *Neural Networks*, 23, 998-1003.
- Viviani, P. (2002). Motor competence in the perception of dynamic events: A tutorial. In W.
  Printz & B. Hommel (Eds.), *Common Mechanisms in Perception and Action* (pp. 406-442). Oxford, UK: Oxford University Press.
- Viviani, P., & Flash, T. (1995). Minimum-jerk, two-thirds power law, and isochrony: Converging approaches to movement planning. *Journal of Experimental Psychology-Human Perception and Performance*, 21, 32-53.
- Volpe, G., D'Ausilio, A., Badino, L., Camurri, A., & Fadiga, L. (2016). Measuring social interaction in music ensembles. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 371(1693). doi:10.1098/rstb.2015.0377
- Vorberg, D., & Hambuch, R. (1978). On the temporal control of rhythmic performance. In J. Requin (Ed.), *Attention and Performance* (Vol. 7, pp. 535-555). Hillsdale, New Jersey: Erlbaum.
- Vorberg, D., & Schulze, H. (2002). Linear Phase-Correction in Synchronization: Predictions, Parameter Estimation, and Simulations. *Journal of Mathematical Psychology*, 46(1), 56-87. doi:10.1006/jmps.2001.1375
- Vorberg, D., & Wing, A. (1996). Modeling variability and dependence in timing. In S. Keele
  & H. Heuer (Eds.), *Handbook of perception and action* (Vol. 2, pp. 181-262). New
  York: Academic Press.

- Wagenmakers, E.-J., Farrell, S., & Ratcliff, R. (2004). Estimation and interpretation of 1/fa noise in human cognition. *Psychonomic Bulletin & Review*, *11*(4), 579-615.
- Walton, A.E., Richardson, M.J., Langland-Hassan, P., & Chemero, A. (2015). Improvisation and the self-organisation of multiple musical bodies. *Frontiers in Psychology*, *6*, 313. Doi: 10.3389/fpsyg.2015.00313.
- Wanderley, M.M. (1999). Non-obvious Performer Gestures in Instrumental Music. In A.A. Braffort, R. Gherbi, S. Gibet, D. Teil, & J. Richardson (Eds.), *Gesture-based Communication in Human-Computer Interaction* (pp. 37-48). Berlin: Springer-Verlag.
- Wanderley, M.M. (2002). Quantitative analysis of non-obvious performer gestures. In I. Wachsmuth & T. Sowa (Eds.), *Gesture and Sign Language in Human-Computer Interaction* (pp. 241-253). Berlin: Springer Verlag.
- Wanderley, M.M., Vines, B.W., Middleton, N., McKay, C., & Hatch, W. (2005). The musical significance of clarinetists' ancillary gestures: An exploration of the field. *Journal of New Music Research*, 34(1), 97-113.
- Wang, C., & Yang, W. (2012). Using detrended fluctuation analysis (DFA) to analyze whether vibratory insoles enhance balance stability for elderly fallers. *Archives of Gerontology and Geriatrics*, 55, 673-676. doi:10.1016/j.archger.2011.11.008.
- Ward, L. M., & Greenwood, P. E. (2007). 1/f noise. Scholarpedia, 2(12), 1537.
- Williamon, A., & Davidson, J. W. (2002). Exploring co-performer communication. *Musicae Scientiae*, 6, 53-72.
- Wilson, C., Simpson, S. E., Van Emmerik, R. E. A., & Hamill, J. (2008). Coordination variability and skill development in expert triple jumpers. *Sports Biomechanics*, 7(1).
- Wing, A. M. (2002). Voluntary timing and brain function: An information processing approach. *Brain and Cognition*, *48*(1), 7-30. doi:10.1006/brcg.2001.1301.

- Wing, A. M. (2014). Coordination in group timing. *International Conference on Timing and Time Perception*, 126, 15-15. doi:10.1016/j.sbspro.2014.02.301.
- Wing, A. M., & Beek, P. J. (2002). Movement timing: a tutorial. Common Mechanisms in Perception and Action, 19, 202-226.
- Wing, A. M., Endo, S., Bradbury, A., & Vorberg, D. (2014). Optimal feedback correction in string quartet synchronization. *Journal of the Royal Society Interface*, *11*(93). doi:ARTN 2013112510.1098/rsif.2013.1125.
- Wing, A. M., Endo, S., Yates, T., & Bradbury, A. (2014). Perception of string quartet synchronization. *Frontiers in Psychology*, 5. doi:ARTN 111510.3389/fpsyg.2014.01115
- Wing, A. M., & Kristofferson, A. (1973). Response Delays and Timing of Discrete Motor Responses. *Perception & Psychophysics*, 14(1), 5-12. doi: 10.3758/Bf03198607.
- Witek, M. A., Clarke, E. F., Wallentin, M., Kringelbach, M. L., & Vuust, P. (2014).
  Syncopation, body-movement and pleasure in groove music. *PLoS One*, *10*(9).
  doi:https://doi.org/10.1371/journal.pone.0094446.
- Witek, M. A., Popescu, T., Clarke, E. F., Hansen, M., Konvalinka, I., Kringelbach, M. L., & Vuust, P. (2017). Syncopation affects free body-movement in musical groove. *Experimental Brain Research*, 235, 995-1005.
- Wöllner, C., & Auhagen, W. (2008). Perceiving conductors' expressive gestures from different visual perspectives. An exploratory continuous response study. *Music Perception: An Interdisciplinary Journal, 26*(2), 129-143. doi:10.1525/mp.2008.26.2.129
- Wöllner, C., & Canal-Bruland, R. (2010). Keeping an eye on the violinist: motor experts show superior timing consistency in a visual perception task. *Psychological Research*, 74(6), 579-585. doi:10.1007/s00426-010-0280-9

- Wöllner, C., Deconinck, J. F., Parkinson, J., Hove, M. J., & Keller, P. E. (2012). The perception of prototypical motion: Synchronization is enhanced with quantitatively morphed gestures of musical conductors. *Journal of Experimental Psychology Human Perception and Performance, 38*, 1390-1403. doi:10.1037/a0028130
- Wolpert, D., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nature Neuroscience*, *3*, 1212-1217. doi:10.1038/81497. PMID 11127840
- Zelaznik, H., Binsted, G., Georgescu, A., & Brownell, K. (2007). Individual differences in timing precision for eye movement timing, tapping and circle drawing: Is saccadic timing shareable with timing in tapping? *Journal of Sport & Exercise Psychology, 29*, S142-S143.
- Zelaznik, H. N., & Rosenbaum, D. A. (2010). Timing Processes Are Correlated When Tasks Share a Salient Event. *Journal of Experimental Psychology-Human Perception and Performance, 36*(6), 1565-1575. doi:10.1037/a0020380
- Zelaznik, H. N., Spencer, R. M., Ivry, R. B., Baria, A., Bloom, M., Dolansky, L., . . .
  Whetter, E. (2005). Timing variability in circle drawing and tapping: probing the relationship between event and emergent timing. *J Mot Behav*, *37*(5), 395-403.
  doi:10.3200/JMBR.37.5.395-403
- Zelaznik, H. N., Spencer, R. M. C., & Ivry, R. B. (2002). Dissociation of explicit and implicit timing in repetitive tapping and drawing movements. *Journal of Experimental Psychology-Human Perception and Performance, 28*(3), 575-588. doi:10.1037//0096-1523.28.3.575
- Zelic, G., Varlet, M., Kim, J., & Davis, C. (2016). Influence of pacer continuity on continuous and discontinuous visuo-motor synchronisation. *Acta Psychologica*, 169, 61-70.

## APPENDIX

## 1. Appendix A. The three pieces of music used as auditory stimuli

Beginning on the following page, PDF files of the scores of the music used as auditory stimuli are appended.