INVESTIGATING TEMPO BIASES IN THE MENTAL REPRESENTATION OF RHYTHM
USING THE METHOD OF SERIAL REPRODUCTION

By

Elisa Kim Fromboluti

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Psychology – Doctor of Philosophy

2016

ABSTRACT

INVESTIGATING TEMPO BIASES IN THE MENTAL REPRESENTATION OF RHYTHM
USING THE METHOD OF SERIAL REPRODUCTION

By

Elisa Kim Fromboluti

This dissertation tested two related hypotheses about the mental representation of rhythms in the sub-second to second time range, namely, a preferred tempo hypothesis and an input noise hypothesis. The preferred tempo hypothesis proposed that the inherent period of an internal oscillatory timekeeper biases the perceptual representation of rhythms, such that perceived rhythm tempo is systematically distorted towards the period of the oscillator. The oscillator period was hypothesized to be about 500 to 600 ms, based on past results showing that the perception and reproduction of sequence tempi are systematically biased towards this tempo. The input noise hypothesis proposed that increased variability in the initial encoding of a rhythm leads to greater bias towards preferred tempo in the perceptual representation of the rhythm.

To test these hypotheses, three experiments used a novel approach to the study of rhythms, which was a serial reproduction (SR) paradigm. In the SR paradigm, participants are exposed to a stimulus and then reproduce it from memory, with their reproduction serving as the stimulus for the next participant in the series. This process is iterated across a series of participants, forming a chain of reproductions in which each participant's reproduction influences the content being transmitted. Typically, the final reproduction in a chain incorporates the accumulation of systematic distortion across reproductions in the chain, thereby magnifying the bias introduced at each iteration. Bias in serial reproductions results from the adjustment of noisy input representations towards a value consistent with prior experience during perceptual

representation. Greater noise in the stimulus input leads to greater bias, and thus faster chain drift.

The experiments herein varied the seed tempo of rhythms initiating SR chains, asked a series of participants to reproduce the timing of iterated rhythms by tapping, and then compared the final reproduced tempo to the seed tempo of the chain. In line with the preferred tempo hypothesis, across experiments, when participants passively observed the stimulus rhythms before reproducing them, chains initiated at tempi ranging from 150 to 1709 ms converged to participants' preferred tempo. In line with the input noise hypothesis, when participants synchronized finger taps with stimulus rhythms during exposure, reproductions were less biased towards preferred tempo, leading chains to converge to preferred tempo more slowly or not at all. A similar pattern of results was obtained regardless of whether SR chains consisted of reproductions by different participants or by the same participant. Also consistent with the input noise hypothesis, when participants were exposed to and reproduced visual rhythms instead of auditory rhythms, reproductions tended to be more biased towards preferred tempo, leading chains to converge more quickly for visual than auditory rhythms. Synchronizing with visual rhythms slowed convergence of SR chains to preferred tempo, but to a lesser degree than for auditory rhythms.

Overall, by testing these hypotheses in the domain of rhythm using the method of serial reproduction, the results of this dissertation advance basic research on tempo biases in rhythm perception and production. More broadly, the findings demonstrate the feasibility of using the serial reproduction paradigm in a novel domain and lay the groundwork for the future application of this approach to probe underlying biases in the mental representation of rhythm.

ACKNOWLEDGEMENTS

clear-thinking and enthusiasm working on several projects in the lab, including components of the ArtPrize project.

I cannot overstate the positive impact that working with the many bright, energetic and creative research assistants in the TAP lab throughout the years has had on me. I would like to extend my thanks to two particular research assistants who facilitated data collection for Experiments 2 and 3 of this dissertation. In addition to collecting data with efficiency and aplomb, Stacey Rowland never failed to relate an interesting anecdote or historical story, and Alyssa McClelland's enthusiasm and creative problem-solving were infectious.

David McFarlane also deserves special thanks for his thorough and tireless assistance with configuring the Arduino-based tapping boards used in Experiments 2 and 3.

I do not think I would have made it through graduate school without the friendship of fellow lab-mates and peers, in particular, Katherine Jones, Lauren Amick, Shuli Yu, Jenn Lampen, Peter Kvam, Chad Peltier, Stella Qian, and the other graduate students in the Cognition and Cognitive Neuroscience group. Your collective expertise, creativity, and motivation made for an enriching work and social environment.

I would also like to thank my family for their support throughout the years. My mom, in particular, has spent many hours acting as a sounding board for research ideas and frustrations.

Finally, I would like to thank my husband, Matthew Fromboluti, for his unflagging support throughout this process! Many a well-timed tasty snack, cat attack, or small draught of wine kept me on track, bolstered my mood, or reminded me to take a break during times when my morale dwindled. Along these lines, Virgil Fromboluti (my feline co-author) also deserves thanks for keeping me company throughout the writing process.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INVESTIGATING TEMPO BIASES IN THE MENTAL REPRESENTATION OF RHYTHM USING THE METHOD OF SERIAL REPRODUCTION

Rhythm, in psychology, refers to the perception of the regular, temporal organization of sequential events (Swindle, 1913). Where we feel the beat in the music we listen to, what beat we dance to, or the words we feel are emphasized in a poem are examples of rhythm perception on the sub-second to second time scale in daily life. The sub-second to second time scale is of particular interest because it encompasses the time range of perceivable rhythms, which are characteristic of patterns in music and speech (Giraud & Poeppel, 2012; Grahn, 2012; Greenberg, Carvey, Hitchcock, & Chang, 2003; Jackendoff & Lerdahl, 1981; London, 2004; McAuley, 2010; Parncutt, 1994; Patel, 2003; Peelle & Davis, 2012; Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995). More broadly, time intervals in this range are on the scale of rhythms in our dynamic world that have been shown guide attention from moment to moment in time (Barnes & Jones, 2000; Jones, 1976; Jones & Boltz, 1989; Jones, Johnston, & Puente, 2006; Jones, Kidd, & Wetzel, 1981; Jones, Moynihan, MacKenzie, & Puente, 2002; Large & Jones, 1999; McAuley & Jones, 2003; J. E. Miller, Carlson, & McAuley, 2012).

One important aspect of rhythm perception is tempo (Baese-Berk et al., 2014; McAuley, 2010; McAuley & Semple, 1999; J. L. Miller, Green, & Schermer, 1984; Morrill, Baese-Berk, Heffner, & Dilley, 2015; Parncutt, 1994; van Noorden & Moelants, 1999). Tempo refers to the rate or pace of events – how quickly or slowly they proceed – which is determined by the duration separating event onsets. A sizeable body of previous literature has consistently found that most adults share a preferred tempo of approximately 500 to 600 ms (Baruch, Panissal-Vieu, & Drake, 2004; Drake, Jones, & Baruch, 2000; Fraisse, 1982; McAuley, Jones, Holub, Johnston,

& Miller, 2006; Mishima, 1956; Parncutt, 1994; Stern, 1900; van Noorden & Moelants, 1999; Vanneste, 2001; Vierordt, 1868; Wallin, 1911; Woodrow, 1951). From an entrainment perspective on timing and temporal attention (Jones, 1976; Large & Jones, 1999; McAuley & Jones, 2003), preferred tempo reflects the inherent period of an internal oscillatory timekeeper or set of oscillations (Drake et al., 2000; McAuley et al., 2006). The period of the internal oscillator transiently adjusts to the tempi of stimuli in the environment, yielding an internal representation of environmental rhythms, but either drifts back to its inherent period in the absence of continued stimulation, or adjusts only partially in response to the stimulus (Large & Kolen, 1994; McAuley, 1995), leading to bias. Although preferred tempo has been shown to bias the production of rhythmic sequences at a wide range of tempi (McAuley et al., 2006) and to influence motor excitability during passive listening (Michaelis, Wiener, & Thompson, 2014), the effect of inherent preferences for tempo on the perception of the temporal structure unfolding around us each moment, such as the structure of rhythmic sequences, is still not fully understood.

This dissertation tested two related hypotheses about the mental representation of rhythms in the sub-second to second time range, namely, a preferred tempo hypothesis and an input noise hypothesis. The preferred tempo hypothesis proposes that the inherent period of an internal oscillatory timekeeper biases the perceptual representation of rhythms, such that perceived rhythm tempo is systematically distorted towards the period of the oscillator. The input noise hypothesis proposes that variability in the initial encoding of a rhythm will influence the amount of bias towards preferred tempo present in perceptual representations, with noisier initial encoding resulting in greater bias towards preferred tempo in perceptual representations.

The preferred tempo and input noise hypotheses were tested in three experiments that used a novel, serial reproduction (SR) paradigm. In general, the SR paradigm that was used in

the experiments is similar to a game of telephone (Bartlett, 1932). Generally in this paradigm, the first participant in a series is presented with a seed stimulus that they attempt to reproduce as accurately as possible. The first participant's reproduction serves as the stimulus for the next participant in the series, and so on, forming a chain of reproductions in which each participant's reproduction influences what is presented to the next participant. The reproduction of the final participant in the chain represents the accumulated distortion of the stimulus across all participants in the chain.

The SR paradigm may be used to reveal psychological biases about a particular domain (e.g., visual magnitude, color or time perception) by varying the value of seeds along a single dimension in that domain (e.g., width, hue or duration) across separate chains and then comparing the value produced by the final participant in all chains or the net drift in reproduced values across iterations. Separate chains initiated at different seed values are predicted to converge by the final generation upon a common value that reflects the underlying bias of participants in the chain. Drift from seed value to final value is expected to be greater for chains initiated by seeds that are more different from the underlying expectation that biases reproductions.

Bias in serial reproduction is introduced as a result of each participant in an SR chain implicitly compensating for noise in the observation of a stimulus by shifting the value of their perceptual representation and subsequent reproduction towards a value consistent with expectations (Griffiths & Kalish, 2005, 2007; Kalish, Griffiths, & Lewandowsky, 2007; Xu & Griffiths, 2010). The implicit goal in performing this automatic adjustment is to maximize the accuracy of the reproduction (Hemmer & Steyvers, 2009; Huttenlocher, Hedges, & Vevea, 2000). The basic logic is that participants implicitly use knowledge of what is reasonable based

on previous lifetime experience in order to reduce the uncertainty associated with a given, noisy input. Participants in an SR chain are presumed to have similar expectations with respect to the domain of the chain, and thus to introduce similar, systematic error with each reproduction. In this way, each participant's reproduction introduces systematic error into the chain that reflects a bias towards expectations. Passing information through the chain magnifies biases by propagating systematic errors across generations, thereby revealing group-level expectations for the domain of the chain.

The experiments in this dissertation used the SR paradigm to examine whether and to what extent inherent tempo biases influence the mental representation of rhythms. In all three experiments, the first participant in an SR chain was presented with a brief, isochronous seed sequence and then was asked to reproduce the timing of the sequence by tapping. The stimulus rhythm that the next participant was presented with was based on the preceding participant's reproduction, just as their reproduction was used to generate the stimulus rhythm for the following participant. The tempi of seed stimuli were manipulated so that separate chains began at tempi ranging from fast to slow (relative to typical preferred tempo). The overarching prediction of the preferred tempo hypothesis across experiments was that SR chains initiated with a variety of seed tempi would converge to a tempo of approximately 500 to 600 ms, reflecting an underlying preference for this tempo (Baruch et al., 2004; Drake et al., 2000; Fraisse, 1982; McAuley et al., 2006; Mishima, 1956; Parncutt, 1994; Stern, 1900; van Noorden & Moelants, 1999; Vanneste, 2001; Vierordt, 1868; Wallin, 1911; Woodrow, 1951).

Experiment 1 tested the preferred tempo hypothesis using seed tempi that ranged from 300 to 1200 ms. Participants passively observed and then reproduced stimulus rhythms in a "portable lab" stationed at a public venue. Regardless of whether the seed tempo initiating an SR

chain was fast or slow, later rhythm reproductions in the chain were predicted to converge to an intermediate tempo of approximately 500 to 600 ms, corresponding to previous observations of preferred tempo. To preview results, tempo reproductions in Experiment 1 were systematically biased towards approximately 500 ms, consistent with predictions. Thus, Experiment 1 established that the method of serial reproduction is a viable tool for uncovering a perceptual bias for tempo.

Experiment 2 tested both preferred tempo and input noise hypotheses in the laboratory using seed tempi ranging from 150 to 1709 ms. Based on the preferred tempo hypothesis, chains initiated by differing seed tempi were predicted to converge to a tempo of approximately 500 to 600 ms, replicating Experiment 1. In addition to the SR task, each participant in Experiment 2 generated several sequences of self-paced taps at a comfortable rate in order to assess spontaneous motor tempo (SMT). SMT is a commonly-used method for assessing the preferred tempo of an individual (Drake et al., 2000; Fraisse, 1982; McAuley et al., 2006; Stern, 1900). For most adults, SMT is typically in the range of 400 to 700 ms (McAuley et al., 2006; Stern, 1900). SMT has been proposed to provide an estimate of the inherent ("preferred") period of an internal oscillatory timekeeper, based on the general consistency of SMT across separate productions by the same individual (Rimoldi, 1951), systematic changes in SMT across the lifespan (McAuley et al., 2006), and correlation with perceptual ratings of preferred tempo (Fraisse, 1982; McAuley et al., 2006). More broadly, SMT has been proposed to reflect the inherent rate of mental activity (Stern, 1900). Thus, in Experiment 2, the preferred tempo hypothesis was further tested by examining whether SR chains converged, specifically, to the average SMT of participants who contributed reproductions.

Experiment 2 also tested the input noise hypothesis, which proposed that variability in the initial encoding of a stimulus rhythm would affect the amount of bias towards preferred tempo present in the perceptual representation of that rhythm. Applied to SR chains, this hypothesis predicts that how quickly (or whether) chains converge to preferred tempo will depend on levels of input noise for stimulus rhythms. Faster convergence is predicted when input noise for the stimulus is high than when input noise is low.

The motivation for considering the influence of input noise on bias in the perceptual representation of rhythms comes from two previous lines of work, one on modality differences in the encoding of rhythms and another on the role of movement during exposure to rhythms in rhythm perception. A large number of studies have found evidence for distinctions in the representation of temporal information between auditory and visual modalities (Becker & Rasmussen, 2007; Collier & Logan, 2000; Grondin & McAuley, 2009; McAuley & Henry, 2010; Recanzone, 2003). Generally, discrimination of duration and tempo changes is superior for auditory compared to visual stimuli (Collier & Logan, 2000; Grondin & McAuley, 2009). Similarly, motor synchronization with or reproduction of rhythms tends to be more accurate and less variable when the stimulus is auditory than when it is visual (Dunlap, 1910; Patel, Iversen, Chen, & Repp, 2005; Repp, 2003, 2005; Repp & Su, 2013). When rhythms are presented in both modalities (either simultaneously or in succession), auditory rhythms potentially prime or drive perceptual interpretations of visual rhythms (Guttman, Gilroy, & Blake, 2005; McAuley & Henry, 2010; Recanzone, 2003) as well as biasing motor synchronization and reproduction (Repp & Penel, 2002, 2004), whereas visual rhythms do not appear to affect perception or reproduction of auditory rhythms.

In Experiment 2, the modality of stimulus rhythms was manipulated in order to test the predictions of the input noise hypothesis and examine differences in input encoding noise across modalities. Different groups of participants were exposed to isochronous stimulus rhythms that consisted of either a sequence of tones (auditory condition, similar to Experiment 1), or a sequence of circles (visual condition). Faster convergence of chains to preferred tempo was predicted for visual compared to auditory chains, based on previous research suggesting that the tempo of visual sequences is less accurately encoded than the tempo of auditory sequences (Grahn, Henry, & McAuley, 2011; Hove, Fairhurst, Kotz, & Keller, 2013; London, Burger, Thompson, & Toiviainen, 2016; McAuley & Henry, 2010; Patel et al., 2005; Repp & Penel, 2002).

As a further test of the input noise hypothesis, and in order to investigate the role of motor activity in the perceptual representation of auditory and visual rhythms, in Experiment 2, the task participants engaged in during exposure to stimulus rhythms was also manipulated. Participants either passively observed the stimulus rhythms during exposure (similar to Experiment 1), or synchronized finger taps with the rhythms during exposure. Several behavioral studies have shown that synchronous movement improves the accuracy of judgments of stimulus timing. For example, Iordanescu, Grabowecky, and Suzuki (2013) found that thresholds for judging the temporal location of a target tone improved (i.e., were smaller) when the onset of the stimulus coincided with a trial-initiating button-press by the participant, compared to passively initiated trials. In a similar vein, Manning and Schutz (2013) found that judgments about whether a probe tone was 'on time' relative to the timing of a preceding isochronous rhythm were more accurate when participants synchronized with the rhythm than when they refrained from moving during rhythm presentation. Greater accuracy in judging probe tone timing when participants

synchronized with the context rhythm was interpreted as the result of improved timing acuity conferred by tapping. Consistent with this interpretation, neurophysiological studies have further found that synchronizing movement with a periodic auditory stimulus enhances cortical and brainstem responses (Nozaradan, Schönwiesner, Caron-Desrochers, & Lehmann, 2016), and that individuals with more consistent synchronization of finger taps with a tone sequence also had more consistent brainstem response to a repeated speech sound (Tierney & Kraus, 2013). Together, these behavioral and neurophysiological results point to the beneficial role of explicit movement in the accurate representation of (rhythmic) temporal information.

More broadly, there has been much recent interest in the potential involvement of the motor system in predictive timing (Arnal, Doelling, & Poeppel, 2014; Arnal & Giraud, 2012; Fujioka, Trainor, Large, & Ross, 2012), and particularly in the potential ability of neural oscillatory activity in frequency bands typically associated with motor activity to modulate sensory inputs on the basis of temporal expectations – sometimes called 'active sensing' (Morillon & Schroeder, 2015; Morillon, Schroeder, & Wyart, 2014; Schroeder, Wilson, Radman, Scharfman, & Lakatos, 2010). The general idea behind active sensing proposals is that sensory and motor systems operate in a coordinated manner to acquire and optimize information intake from the environment. Neural oscillatory activity in frequency bands associated with movement (such as beta oscillations ~20 Hz) aligns with regular (rhythmic) external stimulation as a result of either explicit or 'simulated' movement (Arnal et al., 2014; Morillon, Schroeder, Wyart, & Arnal, 2016). This activity, in turn, entrains activity in sensory areas, effectively modulating the phase and period of lower-level responses such that sensory representations of events occurring at expected time-points (i.e., in phase with the motor representation) are sharpened relative to unexpected (out-of-phase) events. Even in the absence of movement, some form of motor system

involvement in the representation of temporally regular patterns (such as those found in rhythms) has been suggested to facilitate more accurate representation of temporal aspects of those patterns (i.e., reduce encoding noise, Arnal et al., 2014).

Thus, in Experiment 2, synchronizing with the stimulus during exposure was expected to reduce input noise relative to passively observing the stimulus (Manning & Schutz, 2013; Morillon et al., 2014; Nozaradan et al., 2016). Based on this expectation, the input noise hypothesis predicted that there would be reduced bias towards preferred tempo in the synchronize compared to the observe condition, leading chains in the synchronize condition to converge more slowly (if at all) to preferred tempo. Based on previous studies finding reduced synchronization accuracy with visual compared to auditory rhythms (Patel et al., 2005; Repp, 2003; Repp & Penel, 2002, 2004), synchronization during exposure to visual rhythms was hypothesized to potentially reduce encoding noise *less* for visual compared to auditory rhythms. This led to the prediction that reproductions in the visual-synchronize condition might be more biased towards preferred tempo compared to the auditory-synchronize condition, causing chains in the visual-synchronize condition to converge towards preferred tempo more quickly than their auditory counterparts.

Finally, Experiment 3 tested the preferred tempo and input noise hypotheses using single-participant SR chains. As in Experiment 2, input noise was varied by manipulating exposure task (synchronize or observe) across groups of participants, and each participant was initially presented with isochronous auditory sequences at seed tempi ranging from 150 to 1709 ms. The pattern of chain drift in Experiment 3 was predicted to parallel the pattern predicted in Experiment 2. Based on the preferred tempo hypothesis, when participants passively observed the stimulus rhythms, chains were predicted to converge to the preferred tempo of the single

participant who contributed reproductions to the chain (as measured by SMT). Based on the input noise hypothesis, when participants synchronized with the stimulus rhythms, chains were expected to converge more slowly or not at all to preferred tempo. Variation in the pattern of results relative to Experiment 2 was considered in light of individual differences in preferred tempo or tempo limits of the participants.

CHAPTER 2: EXPERIMENT 1. SERIAL REPRODUCTION OF RHYTHMS ACROSS A

RANGE OF TEMPI

Experiment 1 used a serial reproduction (SR) paradigm to test the preferred tempo hypothesis with auditory rhythms. In Experiment 1, the seed tempi of the initial isochronous auditory rhythms in SR chains were varied across individuals. According to the preferred tempo hypothesis, regardless of the seed tempo of the stimulus rhythm initiating a chain of reproductions, chains were predicted to converge to a preferred tempo of approximately 500 to 600 ms (Baruch et al., 2004; Drake et al., 2000; Fraisse, 1982; McAuley et al., 2006; Mishima, 1956; Parncutt, 1994; Stern, 1900; van Noorden & Moelants, 1999; Vanneste, 2001; Vierordt, 1868; Wallin, 1911; Woodrow, 1951). For Experiment 1, the fixed inter-onset interval (IOI) of the isochronous rhythms initiating each chain took on one of four values: 300, 600, 900, or 1200 ms. The first participant in a chain heard one of these isochronous seed rhythms and then reproduced its timing by tapping. The values of the sequence of inter-tap intervals (ITIs) from their reproduction were then used to generate the sequence of IOIs for the stimulus rhythm presented to the next participant in the chain.

Chains initiated with a seed tempo that was faster than preferred tempo (i.e., the 300 ms seed tempo) were predicted to demonstrate positive drift, reflecting gradual slowing down across reproductions towards preferred tempo. In contrast, chains initiated with a seed tempo that was slower than preferred tempo (i.e., the 900 ms and 1200 ms seed tempi) were predicted to demonstrate negative drift, reflecting successive speeding up of across reproductions. The tempi of chains initiated at a value similar to preferred tempo (i.e., the 600 ms seed tempo) were predicted to drift less than the other chains. Thus, across chains in Experiment 1, drift was predicted to have a negative linear relationship with seed tempo (negative slope indicates

convergence of chains). The predicted convergence tempo, which was estimated by calculating the x-intercept of the linear equation for drift, was predicted to be about 500 to 600 ms based on the past literature on preferred tempo.

Data for Experiment 1 were collected in a public setting via a brief interaction with a portable lab, shown in Figure 1. The portable lab consisted of a large metal frame housing an iPad running the experiment. The portable lab was stationed at a venue during ArtPrize7 between September 19th and October 11th, 2015. Any passer-by could participate by following the instructions on the custom iPad app running the experiment.

Method

Design. Seed tempo was a between-subjects factor with four levels. Seed tempo was defined by the fixed inter-onset interval (IOI) of the isochronous auditory rhythm initiating each serial reproduction chain. From fastest to slowest, the IOIs for the four seed tempi were 300, 600, 900, and 1200 ms.

Participants. One hundred eighty-four attendees at ArtPrize7 contributed data to the experiment between September 23rd and October 11th, 2015. Participants were distributed across thirteen chains, each consisting of approximately twelve reproductions (each made by a separate participant). Due to a programming error, one chain had only six participants, and several chains included more than twelve participants (the longest was eighteen). Participants were randomly assigned to chains. There were three, six, three, and one chain(s) initiated by each of the 300, 600, 900, and 1200 ms seed tempi, respectively.

Participant demographic data and musical experience are summarized in Table 1. Age-range categories instead of free-form numerical responses (i.e., "How old are you?") were used in order to encourage responding to this question. A study by Peterson (1984) found that survey

respondents were more likely to skip a question about age when phrased in terms of years or date of birth than when phrased as a categorical choice. The musical experience question was drawn from the Ollen Musical Sophistication Index (OMSI, Ollen 2006). The distributions of age, sex, and musical experience were similar across the four seed conditions.
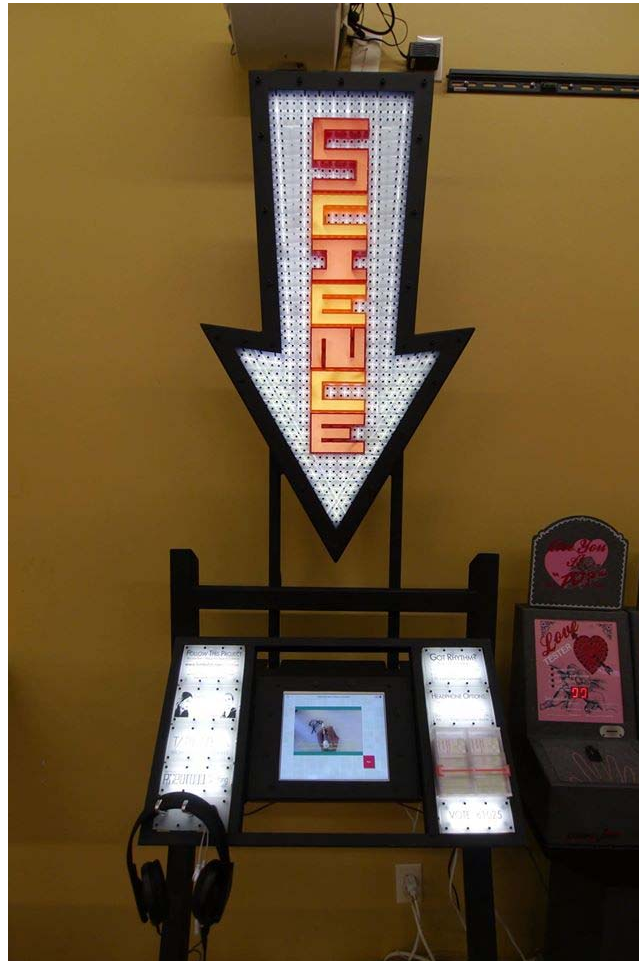
Figure 1. Portable lab used in Experiment 1
The portable lab consisted of a large metal frame housing an iPad running the experiment. The lab was stationed at a public venue during ArtPrize7, an annual art competition in Grand Rapids, MI, between September 19th and October 11th, 2015.

Table 1. Experiment 1: Demographic & musical experience summary by seed tempo

	Seed Tempo				_Total_
	300	600	900	1200	
_N_ Chains	3	6	3	1	13
_N_ Participants	47	78	41	18	184
Sex					
Male	26	38	22	10	84
Female	20	37	19	8	96
Age					
< 17	7	13	8	1	29
18-24	12	22	9	2	45
25-34	11	13	7	9	50
35-44	3	11	6	3	23
45-54	7	8	6	2	23
55-64	4	10	2	1	17
> 65	3	1	3	0	7
Musical Experience					
Non-musician	15	12	10	5	42
Music-loving non-musician	16	31	11	8	66
Amateur musician	11	23	12	2	48
Serious amateur musician	3	9	4	1	17
Semi-professional musician	1	0	1	2	4
Professional musician	1	2	2	0	5
No response	0	1	1	0	2

Note. Values shown are counts by category. If the sum of counts in a category does not match by-condition or total counts, that indicates that some participants chose not to respond to that question. For Age and Musical Experience, the categories shown correspond to response options from the survey participants completed.

Equipment. Data were collected using a portable lab as pictured in Figure 1, which was stationed at a venue during ArtPrize7 in Grand Rapids, MI between September 23rd and October 11th, 2015. The lab was a large metal stand with signage indicating that people could participate in an experiment using the installed iPad. The iPad ran a custom app that presented stimuli and collected data. Auditory stimuli were presented over Sennheiser HD 202 II Professional

Headphones. Participants were able to adjust the volume to a comfortable listening level before starting the experiment.

Stimuli. Stimulus rhythms were sequences of twelve 50 ms sine tones presented at 400 Hz. The duration between tone onsets was varied to create the different seed tempi, as shown in Figure 2. Seed sequences consisted of twelve repeated tones separated by one of four possible fixed IOIs, corresponding to different seed tempi. In order of fastest to slowest, these IOIs were 300, 600, 900, and 1200 ms. The first participant in a chain heard one of the isochronous seed sequences. The inter-tap intervals (ITIs) produced by a participant were used as the IOIs for the stimulus presented to the next participant in the chain. Stimuli were generated by the custom iPad app that ran the experiment.
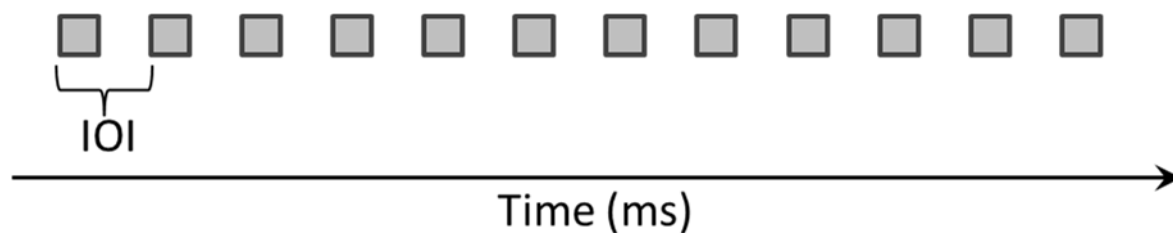
Figure 2. Stimulus diagram
Across experiments, stimulus rhythms were sequences of twelve to thirty-one repetitions of a 50 ms event. In auditory conditions, the event was a 400 Hz sine tone. In visual conditions (Experiment 2 only), the event was a green circle. The duration between event onsets (i.e., the inter-onset interval or IOI) was varied to create different seed tempi, which were presented to the first participant in a chain. After the first participant, stimulus IOIs were based on either the sequence of produced inter-tap intervals (ITIs, in Experiment 1) or the median produced ITI (Experiment 2 and 3) of the preceding reproduction.

Procedure. Any passer-by to the portable lab could participate in the experiment by walking up to the lab and following the instructions provided by the app. A participant initiated the experiment by pressing a button labeled *start* on the iPad screen. The app first presented a practice tapping trial in which the participant tapped twelve times at a steady comfortable pace

on a button labeled "TAP" on the screen of the device. The practice could be repeated until the participant felt comfortable tapping on the screen (if the participant opted to repeat the tapping practice, only data from the final repetition were recorded).

Next, the participant was given brief auditory and visual instructions about the experimental task. During the experimental trial, the participant listened to a twelve-tone stimulus rhythm and then reproduced it by tapping twelve times on the screen of the iPad (yielding eleven ITIs). There were two identical repetitions of the experimental trial. After the first trial ended, the participant pressed a button when ready to proceed with the second trial. For the purposes of stimulus generation and analysis, the first experimental trial was treated as practice for the SR task. Only the ITIs from the second experimental trial were used to generate stimuli for the next participant.

After the experimental trials, participants indicated whether or not they wanted to submit their data. If they chose to submit data, they completed a five-question survey about basic demographic info (zip code, gender, age range, musical experience, other comments). At the end, participants were shown their performance relative to the target rhythm they were trying to reproduce and relative to the seed rhythm for the SR chain to which they belonged. The entire interaction took less than five minutes.

Data Analysis

Preprocessing of tapping data. Preliminary processing of tapping data was performed by the iPad app during the experiment in order to generate stimuli for the following participant. Any tap below 50 ms (i.e., a "bounce") was filtered (excluded) during recording by the app. The app considered the next tap after this bounce as the intended next tap. Any additional pre-processing or exclusion in subsequent analysis (when applicable) is described by section.

Spontaneous motor tempo (SMT). Data from the practice tapping task at the beginning of the experiment were used as a rough measurement of spontaneous motor tempo (SMT; McAuley et al., 2006). SMT reflects a participant's preferred movement tempo, which, within an entrainment framework, could reflect the inherent period of an internal oscillator, and as such serves as an estimate of preferred tempo. Although in the present experiment, framing SMT as a practice made it an imperfect measure of SMT (since some participants made long pauses between taps as they learned to use the device), it was considered to, nonetheless, yield a useful approximation of the preferred tempi of participants.

To clean up the SMT data, ITIs greater than 2000 ms were excluded, resulting in the exclusion of 33 ITIs (out of 2170 total, or about 1.5%). Following the analysis method of McAuley et al. (2006), SMT was then calculated by finding the median of the remaining ITIs produced during the practice trial. Medians rather than means were used because medians are less influenced by extreme outliers (that might result, for example, from pauses or missed taps).

Serial reproduction (SR) chains. No filtering of raw tapping data was performed beyond the preprocessing performed during the experiment by the iPad. The tempo of each participant's reproduction was calculated by finding the median of the eleven ITIs of their response. The rationale for using the median rather than the mean of the ITIs was to diminish the influence of potential outliers, similar to the rationale for this choice in the SMT analysis.

Drift of serial reproductions. Tempo drift in SR chains was calculated by finding the difference between the final reproduced tempo in a chain and the seed tempo of the chain. The magnitude of drift is proportional to the magnitude of tempo change across iterations in an SR chain: a larger value of absolute drift indicates a greater change from seed tempo to final tempo than a smaller value. The sign of drift indicates whether tempo sped up or slowed down over

iterations. Positive drift is obtained when the final tempo is slower than the seed, indicating that the tempo slowed down over iterations; negative drift is obtained when the final tempo is faster than the seed tempo, indicating that the tempo sped up over iterations.

Linear regression was used to fit drift as a function of seed tempo. The sign of the slope of the best fitting line indicated whether chains converged (negative), remained at a constant tempo (zero), or diverged (positive). As chains were expected to converge, the slope was expected to be negative. For a negative slope, the magnitude of the slope indicated how strongly SR chains were biased towards a central value, and as such reflects the convergence rate of chains towards a specific tempo. The x-intercept of the linear equation indicates the predicted convergence tempo of chains (i.e., the seed tempo for which no drift is expected). This provided an estimate of the preferred tempo expressed by the chains.

Results

Spontaneous motor tempo (SMT). SMT has been reported to vary with age or across individuals (Drake et al., 2000; McAuley et al., 2006). Given the diverse age range of the sample in this experiment, some deviation of the observed SMT relative to the canonical value for preferred tempo of 500 to 600 ms was expected. However, SMT was not expected to vary systematically across iterations. Consistent with this expectation, the slope of the linear regression predicting SMT by iteration was not significantly different from zero, $SMT = -2.5*Iteration + 434.2$, $p = .3$, $R^2 = .006$. The average SMT across participants was 414.3 ($SD = 147.4$). This value of SMT was reasonably close to previous observations of SMT (McAuley et al. 2006; Baruch ref; Drake Jones Baruch, 2000), particularly for the age range of this sample, which included several participants below the age of 17. Typically, children have been shown to demonstrate faster on-average SMT than adults (Drake et al., 2000; McAuley et al., 2000).

Serial reproduction. Figure 3A shows data for the thirteen chains (individual lines) initiated by the four seed tempi (300, 600, 900, and 1200 ms). Each point reflects the tempo of a single participant's reproduction. The first participant in a chain heard an isochronous stimulus rhythm presented at one of the seed tempi, and then reproduced the stimulus rhythm from memory, and so on. In this figure, the seed tempo corresponds to iteration zero. Overall, the general pattern of tempo drift across chains initiated by different seed tempi was consistent with the preferred tempo hypothesis. The three chains initiated by the 300 ms seed tempo, which was faster than preferred tempo, slowed down by the final iteration to an average tempo of 537.1 ms ($SD = 117.7$). In contrast, chains initiated by seed tempi that were slower than preferred tempo sped up over generations. The three chains initiated by the 900 ms seed tempo sped up on average to a tempo of 375.1 ms ($SD = 220.5$). Similarly, the single chain initiated by the 1200 ms seed tempo also sped up to a tempo of 275.2 ms. The six chains initiated by the 600 ms seed tempo changed the least across reproductions, speeding up slightly to 489.1 ms ($SD = 104.4$).

To quantify the drift in tempo across reproductions, the difference between the final reproduced tempo and the seed tempo for each chain was calculated. Figure 3B shows drift as a function of seed tempo. Positive drift indicates that the final tempo was slower than the seed, whereas negative drift indicates that the final tempo was faster than the seed. As predicted, there was a systematic slowing of the chains initiated by the fast seed tempo (i.e., 300 ms) and speeding of the chains initiated by the slow slow tempi (i.e., 900 and 1200 ms). Drift was smallest in the chains initiated at the tempo closest to previous observations of preferred tempo (i.e., the 600 ms seed tempo), and increased linearly in magnitude for chains initiated by faster or slower seed tempi. The finding that the magnitude of reproduction drift increased as the seed tempo became more distant (either faster or slower) from preferred tempo is consistent with a

description of reconstruction from memory in which extreme stimulus values undergo greater adjustment towards the central value of the prior than less extreme stimulus values (Huttenlocher, Hedges, & Vevea, 2000).
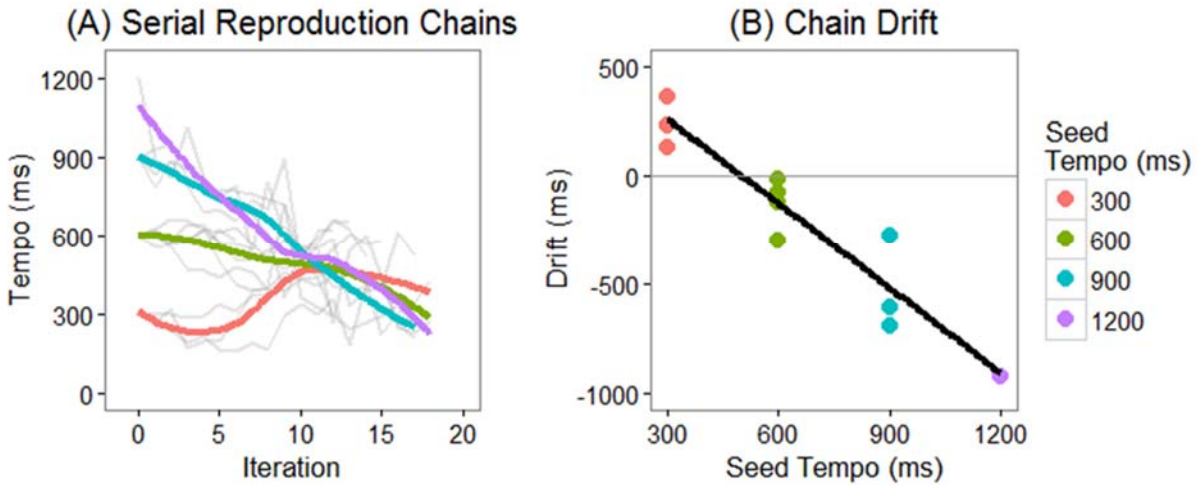
Figure 3. Serial reproduction chains in Experiment 1
Tempo reproduction by participant (light grey lines, Panel A) at each iteration for each chain. Colored lines show trend by seed tempo. Panel B shows chain drift as a function of seed tempo. Each point represents drift for a single SR chain. The black line shows the best-fit line through the data (the drift function). The x-intercept of the drift function estimates the convergence tempo of the chains.

Next, to assess the convergence of chains initiated by different seed tempi, a linear regression line was fit to drift as a function of seed tempo. Consistent with the preferred tempo hypothesis, the slope of the drift function was negative and significantly different from zero, *Drift* = -1.3*Tempo + 646.4, $p < .001$, $R^2 = .88$, indicating that chains converged. The estimated convergence tempo of the chains was 500.1 ms, which was also consistent with the preferred tempo hypothesis and with previous observations of preferred tempo.

Discussion

The preferred tempo hypothesis proposed that the representation of rhythm is systematically biased towards a preferred tempo of about 500 to 600 ms, based on previous research that has found that most adults have a preferred tempo in this range (Baruch et al., 2004; Drake et al., 2000; Fraisse, 1982; McAuley et al., 2006; Mishima, 1956; Parncutt, 1994; Stern, 1900; van Noorden & Moelants, 1999; Vanneste, 2001; Vierordt, 1868; Wallin, 1911; Woodrow, 1951). To test this hypothesis in this experiment, the seed tempi initiating SR chains were manipulated across chains in which participants heard a stimulus rhythm and then reproduced it by tapping. Inter-tap intervals from one participant's reproduction were used as the inter-onset intervals of the stimulus rhythm for the next participant in the chain. Based on the preferred tempo hypothesis, chains were predicted to converge to preferred tempo, regardless of the seed tempo initiating the chain, thereby revealing an underlying preference for a tempo of about 500 to 600 ms.

Consistent with the predictions of the preferred tempo hypothesis, chains initiated by a seed tempo that was faster than preferred tempo (i.e., the 300 ms seed) slowed down over reproductions. In contrast, chains initiated by a seed tempo that was slower than preferred tempo (i.e., the 900 ms and 1200 ms initiated chains) sped up. The tempi of chains initiated at a value similar to preferred tempo (i.e., the 600 ms seed) changed the least. The change in tempo from seed to final generation was formalized as *drift* (the signed difference between these two tempi). Positive drift indicated that the final tempo was slower than the seed, whereas negative drift indicated that the final tempo was faster than the seed. As predicted, fitting a linear function to drift as a function of seed tempo indicated that chains converged (as evidenced by a negative, non-zero slope). Moreover, the predicted convergence tempo was 500.1 ms, which was

21

consistent with previous observations of preferred tempo, and similar to the average SMT of the sample, which was 414.3 ms.

From an entrainment perspective, preferred tempo has been proposed to reflect the inherent period of an internal oscillatory timekeeper (McAuley et al., 2006). From this perspective, reproduction bias for each iteration in a chain reflects the incomplete period adjustment of a participant's internal oscillatory timekeeper to the tempo of the stimulus rhythm. Convergence of SR chains initiated at a variety of seed tempi to preferred tempo, then, can be understood as the accumulation of bias towards preferred tempo, and eventual settling to an equilibrium state, which reflects the inherent period of the internal oscillator. In contrast, convergence of chains to a preferred tempo is generally inconsistent with an interval perspective on timing (Church, Miller, Meck, & Gibbon, 1991; Gibbon, 1977; Gibbon, Church, & Meck, 1984; Treisman, 1963), which would predict no systematic bias in tempo reproductions (Drake & Botte, 1993; Wing & Kristofferson, 1973a, 1973b), and thus no drift across serial reproductions.

Overall, the results that SR chains converged to a tempo of about 500 ms regardless of seed tempo and that drift was a linear function of seed tempo were consistent with both the specific prediction of the preferred tempo hypothesis, and more broadly with the description of bias in serial reproduction as implicit adjustment during perceptual representation of noisy inputs towards a value consistent with expectations (Griffiths & Kalish, 2005, 2007; Huttenlocher et al., 2000; Kalish et al., 2007; Xu & Griffiths, 2010). According to this description, the weighting between stimulus inputs (recent memory for stimulus tempo) and preferred tempo (the expectation) is determined by the ratio of noise (or uncertainty) in each. Extended to SR chains, the noise ratio determines the convergence rate of chains (Xu & Griffiths, 2010). When input

noise associated with a stimulus is large, chains will converge more quickly to the mean of the underlying representation that biases reproductions (i.e., preferred tempo, in this case) than when input noise is small. This property of the SR paradigm offers an opportunity to explore other issues in the representation of rhythm, such as encoding differences across modalities (Recanzone, 2003; Welch & Warren, 1980) and the contribution of movement to the representation of rhythm (Doelling, Arnal, Ghitza, & Poeppel, 2014; Fujioka et al., 2012; Schroeder et al., 2010), as well as differences in encoding across tempi, which will be pursued in Experiment 2.

One outstanding issue from Experiment 1 concerns the relation between the obtained measure of spontaneous motor tempo (SMT) and tempo convergence in the SR chains. Generally, SMT is measured by asking participants to tap at a comfortable, steady rate and is taken as a proxy measure of preferred tempo (Drake et al., 2000; McAuley et al., 2006). Past observations for SMT range from as fast as 200 to as slow as 1600 ms (Drake et al., 2000; McAuley et al., 2006; McAuley, 2010), with the most common or representative value being about 500 to 600 ms. In the present experiment, the average SMT for the sample of 414.3 ms was slightly faster than the representative value of SMT. There are two potential explanations for this discrepancy. First, the ages of the participants in this experiment covered a wide range, as shown in Table 1. McAuley et al. (2006) found that the average SMTs for both children aged 10-12 years old (average SMT = 549 ms) and adults aged 39-59 years old (average SMT = 522 ms) were faster compared to that of adults aged 18-38 years old (average SMT = 630 ms). Younger children aged 4 to 10 years old were found to have SMTs as fast as 300 to 500 ms. The present experiment included both many younger people below 17 years old (29 participants), and several adults in the 39-50 year old range (approximately 50 participants). Thus, it is possible

that the diversity of the age of participants contributed to the SMT being faster than previous observations. Nonetheless, the SMT observed was still in line with the range of values of preferred tempo reported previously in the literature (Drake et al., 2000; Fraisse, 1982; McAuley, 2010; McAuley et al., 2006).

The second potential cause for SMT to have deviated from expectations might have had to do with framing the task as practice. SMT data were obtained from the practice tapping trial that was presented before the experimental trials. Although participants were instructed to tap at a comfortable rate on the practice trial (consistent with the usual instructions given when obtaining SMT data), they were also told that they were tapping for practice. I observed several participants take breaks between taps during the practice trial to ask questions or re-read instructions, suggesting that the method used may not have been an entirely reliable way to assess SMT. In addition, past studies have asked participants to produce longer sequences (e.g., McAuley et al., 2006 had participants tap thirty-one times, and repeated this in multiple trials), whereas in this experiment, SMT calculations were based on only eleven ITIs.

An additional concern related to SMT is that – although the convergence tempo of SR chains was close to expectations based on previous literature – it did not match the preferred tempo, as measured by SMT, of this group of participants. Concerns regarding the reliability of SMT in the present experiment and the relationship to convergence tempo of SR chains were addressed in Experiment 2 by improving the measurement of SMT. Experiment 2 used the same task as in Experiment 1 to assess SMT, but changed the instructions (so participants would know that their goal was to tap at a comfortable rate and to produce an isochronous sequence), and took multiple measurements, each with a greater number of taps than were made during the practice portion of Experiment 1.

Conclusion

In Experiment 1, serial reproduction chains initiated at seed tempi spanning 300 to 1200 ms converged to a tempo of about 500 ms. Overall, this result supports the preferred tempo hypothesis and establishes that the serial reproduction method is a viable tool for examining biases in the mental representation of rhythm. One question raised by this experiment is whether there are predictable conditions under which preferred tempo does not bias the representation of rhythm. This question was motivated by consideration of a description of bias in serial reproduction as the result of adjusting stimulus inputs towards a value consistent with expectations during perceptual representation (Huttenlocher, Hedges, and Vevea, 2000; Xu & Griffiths, 2010). According to this description, the ratio of noise associated with a stimulus input to the noise or uncertainty in the underlying representation that biases reproductions determines the convergence rate of SR chains. This property of SR chains offers an opportunity to explore other issues in the representation of rhythm, such as the role of rhythm modality and the contribution of movement to the representation of rhythm, as well as encoding differences across tempi. This possibility is explored in Experiment 2.

CHAPTER 3: EXPERIMENT 2. THE EFFECT OF INPUT NOISE ON SERIAL REPRODUCTION OF AUDITORY AND VISUAL RHYTHMS ACROSS A BROAD RANGE OF TEMPI

There were two primary goals of Experiment 2. The first goal was to replicate Experiment 1 in the laboratory using a broader range of tempi. Experiment 2 extended the range of tempi initiating serial reproduction (SR) chains by having each participant provide reproductions for seven chains, each initiated at a different seed tempo that ranged from 150 to 1709 ms, in equal log-spaced steps. Participants were grouped into cohorts (as shown in Figure 4), such that the tempo reproductions from one participant for all seven chains were used to generate isochronous stimuli for the next participant in the cohort.

As in Experiment 1, the preferred tempo hypothesis predicted that SR chains would converge to the average preferred tempo of the participants in the chains. In order to evaluate whether the average preferred tempo of the participants in this experiment was consistent with the preferred tempo reported in the literature of approximately 500 to 600 ms, spontaneous motor tempo (SMT) data for each participant were also collected. For the SMT task, participants tapped thirty-one times in a steady rhythm at a tempo they considered comfortable. Although SMT is a measure of preferred *production* tempo, McAuley et al. (2006) found that SMT is very highly correlated with preferred *perceptual* tempo, suggesting that both may share a common underlying mechanism (i.e., from an entrainment perspective, the inherent period of an internal oscillatory timekeeper).

The second goal of Experiment 2 was to test an input noise hypothesis. This hypothesis proposed that variability in the initial encoding of a stimulus rhythm would affect the amount of bias towards preferred tempo present in the perceptual representation of that rhythm. When input

noise is high, chains converge more quickly to expectations (preferred tempo, in this case) than when input noise is low. Applied to the domain of rhythm, greater input noise for the tempo of a stimulus rhythm was hypothesized to lead the perceptual representation of tempo to rely more on preferred tempo (i.e., to be more biased) than conditions of reduced input noise, thereby causing faster convergence of serial reproductions toward an underlying preferred tempo. In order to manipulate input noise, both the modality (auditory vs. visual) of the stimulus rhythms and the participant's task during stimulus exposure (synchronize vs. observed) were varied.

Based on previous literature, visual rhythms were hypothesized to have greater input noise than auditory rhythms. Supporting the view that there should be greater input noise for visual rhythms than auditory rhythms, previous studies have shown that representation of temporal information is less precise for visual rhythms than for auditory rhythms in both perception (Collier & Logan, 2000; Guttman et al., 2005; London et al., 2016; McAuley & Henry, 2010) and production (Dunlap, 1910; Patel et al., 2005; Repp, 2003; Repp & Penel, 2002). For example, one study by McAuley and Henry (2010) found that participants were both more sensitive to tempo changes and were also more likely to perceive an implied beat in auditory compared to visual rhythms. In addition, for an implied beat-period of 600 ms, the perception of a beat in visual rhythms could be primed by prior presentation of auditory rhythms with identical temporal structure, but the reserve was not true.

Parallel findings in studies of rhythm production suggest that encoding of the temporal properties of auditory rhythms may be more precise than it is for visual rhythms. Participants have been shown to be capable of synchronizing movements (finger taps) with successive event onsets separated by as little as 100-120 ms in auditory rhythms, whereas for visual rhythms, as much as 400 ms between successive onsets is required for participants to synchronize

27

movements successfully (Repp, 2003). In one study comparing sensorimotor synchronization to isochronous and non-isochronous auditory and visual rhythms, participants were able to synchronize with isochronous and non-isochronous auditory sequences across a wide range of tempi, but showed generally poor synchronization performance for visual rhythms, particularly when the rhythms were non-isochronous or when synchronization was at fast tempi (Patel et al., 2005). Thus in the present experiment, presenting stimulus rhythms in the visual modality was predicted to increase input noise relative to the auditory modality, leading to faster convergence of chains to preferred tempo compared to auditory chains.

A second factor that was expected to influence input noise was participants' task during exposure to the rhythm. In particular, synchronizing (via finger taps) with the stimulus rhythm during exposure was expected to reduce input noise relative to passively observing the stimulus (Manning & Schutz, 2013; Morillon et al., 2014; Nozaradan et al., 2016), suggesting that there would be reduced tempo drift when participants were asked to synchronize with the stimulus rhythm compared to passive exposure. Results from a recent study provide some neural support for the proposition that synchronizing movement with auditory rhythms improves sensory encoding fidelity. Nozaradan et al. (2016) found that both auditory brainstem and cortical electroencephalographic (EEG) responses to the regular beat of an isochronous auditory stimulus were enhanced (specifically, larger-amplitude steady-state evoked potentials were observed) when participants synchronized finger taps with alternating beats of the stimulus relative to when participants passively listened to the same stimulus. Enhanced response at the brainstem level was interpreted as evidence of modulation of the sensory representation of stimulus timing by synchronous movement (presumed to correspond to decreased variability in the representation of the stimulus). This result was consistent with earlier results finding a systematic relationship

between brainstem response and synchronization performance (Tierney & Kraus, 2013) – i.e., participants who demonstrated a larger amplitude response to a repeated speech sound also tended to synchronize better with an isochronous tone sequence – as well as with the proposition that synchronized movement improves higher-level auditory representations by increasing the gain in the sensory signal (Morillon et al., 2014; Schroeder et al., 2010).

In the present study, participants in the synchronize condition synchronized finger taps with the stimulus rhythm during exposure, whereas participants in the observe condition passively observed the stimulus rhythm during exposure. In both conditions, participants reproduced the stimulus following the exposure phase. Synchronizing with the stimulus during the exposure phase was presumed to reduce input noise associated with the stimulus (Manning & Schutz, 2013; Nozaradan et al., 2016) relative to passively observing the stimulus. During perceptual representation and reproduction, participants who previously synchronized with the stimulus rhythm were hypothesized to rely to a greater degree on this less variable stimulus memory than prior tempo preferences, in contrast to the participants who previously observed the stimulus, who were hypothesized to rely to a greater degree on prior tempo preferences since the stimulus memory was less reliable. Thus SR chains in the observe condition were predicted to be more likely to converge to preferred tempo than chains in the synchronize condition. Chains in the synchronize condition were predicted to either remain (on average) at the seed tempo across iterations (reflecting greater reliance on stimulus memory by participants in this condition), or to drift towards preferred tempo more slowly than chains in the observe condition (reflecting reduced reliance on preferred tempo relative to the observe condition).

Based on previous research showing that movement synchronization is less precise with visual rhythms than with auditory rhythms (e.g., as briefly reviewed above -- Patel et al., 2005;

Repp, 2003), one additional possibility was that synchronization benefits (in the form of reduced drift) might be unique to auditory sequences. Although synchronizing with auditory rhythms was predicted to reduce chain drift relative to passively observing the rhythms, synchronizing with visual rhythms might not have the same effect. In this case, chains of visual rhythms were predicted to drift towards preferred tempo, regardless of whether participants synchronized with or observed the visual rhythms.

In addition to the SMT task, participants in Experiment 2 performed two additional, self-paced tapping tasks, which were fast and slow tempo limit tasks. For these tasks, participants tapped as quickly or as slowly as possible, while maintaining a steady sense of rhythm, providing measures of the lower and upper tempo limits of a participants' rhythm production ability. These measures were of interest because performance on these tasks has previously been proposed to reflect constraints of an individual's internal oscillatory timekeeper (McAuley et al., 2006). Specifically, the range of tempi produced on the fast and slow limits tasks have been proposed to reflect the range of stimulus tempi that afford stable period adjustment by the internal oscillator (i.e., its entrainment region). Outside of this range, the oscillator can transiently adjust its period to match that of stimulus tempi, but will drift back (i.e., decay) more quickly towards its inherent (preferred) period than for stimulus tempi within the range bounded by the fast and slow tempo limits. Consistent with this proposal, McAuley et al. (2006) found that performance on the fast and slow tempo limits tasks reliably predicted the accuracy of tempo reproductions across a broad range of tempi during a synchronize-continue tapping task. Reproductions of stimulus tempi faster/slower than the fast/slow tempo limits were systematically slowed/speeded towards preferred tempo.

In order to begin to explore the nature of the relationship between fast and slow tempo limits and tempo bias in the SR task, these measures in the present experiment were used to calculate normalized produced range (NPR, McAuley et al., 2006), which is the difference between the two limits, divided by SMT for each participant. Considering the concept of an entrainment region (as measured by NPR) in the present experiment leads to the prediction that bias (and thus drift) may be reduced within a bounded range of tempi centered on preferred tempo. Based on the possibility that input noise for stimulus tempo within an entrainment region would be reduced relative to input noise for stimulus tempi outside of an entrainment region, one outcome that was considered was that chains would converge more quickly when the participants in the cohort had, on average, a narrower NPR (indicating a narrow band of tempi affording stable entrainment) compared to chains generated by cohorts with a wider average NPR.

Method

Design. The experiment implemented a 2 (Modality: auditory, visual) x 2 (Exposure task: synchronize, observe) x 7 (Seed tempo: 150, 225, 337, 506, 759, 1139, and 1709 ms) mixed factorial design. Modality and exposure type were between-subjects factors. Seed tempo was a within-subject factor. The values for the IOIs for the seed tempi were selected based on those used in a previous study (McAuley et al., 2006), and were logarithmically-spaced to sample from a broad range of tempi.

Participants. Three-hundred individuals ($N = 205$ female) from the Michigan State University community participated in the experiment either for course credit or for pay at a rate of \$10/hour. The mean age of participants was 23.1 years (range 17 – 63). For the participants with musical training ($N = 169$), the number of years of training ranged from 0.5 to 29 years ($M = 5.5$ years).

The sample size was based on the goal of constructing five cohorts in each modality x exposure condition, each with a sufficient number of participants for chain convergence to occur. The results of Experiment 1 found that chains generally converged after twelve to eighteen iterations. Based on this finding, the present experiment obtained fifteen reproductions in each chain. Generally, it is desirable to have multiple cohorts of chains in order to allow for comparison between cohorts, in case the behavior of a single cohort is idiosyncratic (Kalish et al., 2007).

The three-hundred participants were distributed across the four experimental conditions (modality x exposure), with seventy-five participants in each condition. Each condition was divided into five "cohorts" as shown in the diagram in Figure 4, with fifteen participants in each cohort. Each cohort consisted of seven chains, each initiated by one of the seven seed tempi. Reproductions by one participant for each of the seven chains initiated by the seven seed tempi were transmitted to the next participant in the cohort (by using the tempo of their reproductions to generate isochronous stimulus sequences for the next participant). Thus each participant provided a reproduction for seven chains. There were 140 SR chains in total, evenly distributed across the seven seeds, five cohorts, and four experimental conditions, with each chain consisting of fifteen iterations.
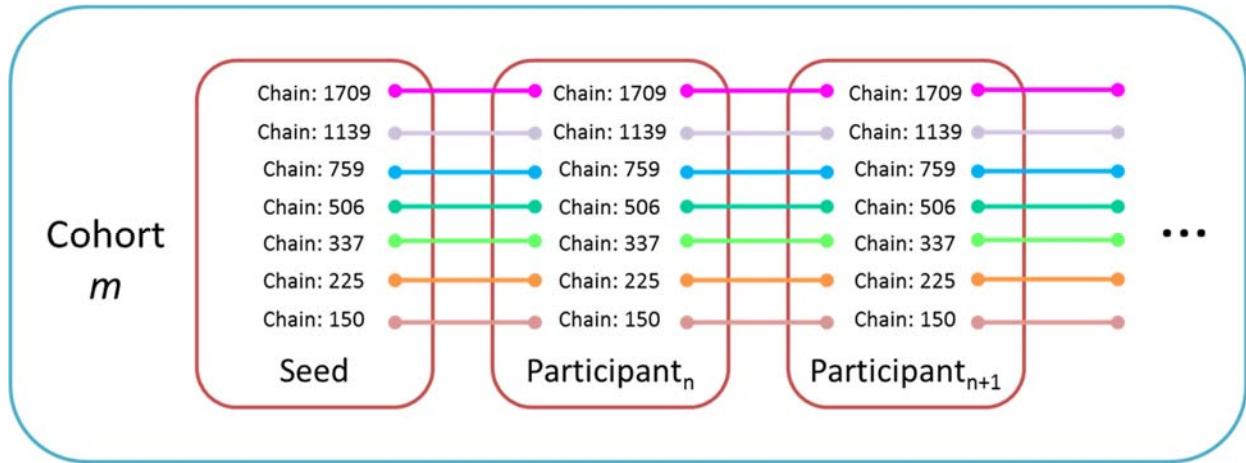
Figure 4. Cohort diagram for Experiment 2
Participants in each condition were grouped into five cohorts. The first participant in a cohort heard or saw isochronous stimulus rhythms presented at the seed tempi and provided seven reproductions (one for each seed tempo). The tempi of their reproductions were used to generate isochronous stimulus rhythms for the next participant in the cohort, and so on. Each cohort consisted of fifteen participants.

Equipment. Stimulus presentation and data collection were performed using a desktop computer running E-Prime. Auditory stimuli were presented over Sennheiser HD 280 Professional Headphones with the volume set to a comfortable listening level. Tapping responses were collected using a custom-built response device consisting of a piezo element affixed to a firm padded surface connected to an Arduino micro-controller. Data collection parameters (number of taps and timeout for receiving responses) were set in E-Prime and sent to the Arduino. On each trial, the Arduino recorded timestamps for taps detected on the piezo (up to the number specified by the data collection parameters or until the end of the timeout interval), and then sent this data to E-Prime for logging.

Stimuli. Both auditory and visual stimulus rhythms consisted of thirty-one repetitions of an event (either a tone or an image of a filled green circle, respectively), as shown in Figure 2 of Experiment 1. The decision to lengthen the stimulus sequences in Experiment 2 compared to Experiment 1 was based on previous use of a similar number of events in a similar tapping

33

paradigm (McAuley et al., 2006; Wing & Kristofferson, 1973), and was intended to provide ample opportunity for participants to learn the tempo of a stimulus sequence. Auditory rhythms were similar to Experiment 1 and consisted of a 400 Hz sine tone presented for 50 ms, repeated thirty-one times, with a fixed IOI. Auditory sequences were generated in MATLAB. Visual sequences consisted of an image of a green circle (on a white background) presented for 50 ms, repeated thirty-one times. The timing of onset and offset of the green circles for the visual sequences was controlled by E-Prime by reading-in target inter-stimulus intervals (ISIs) from a text file for each participant for each chain.

The first generation of participants heard or saw the isochronous seed rhythms. Stimulus generation between participants differed from Experiment 1, in which the ITIs of one participant's reproduction were used as the IOIs for the next participant's stimulus rhythm. For each chain in Experiment 2, the mean *tempo* of one participant's ITIs was used to generate an isochronous stimulus sequence for the next participant in the cohort. Specifying the timing of stimuli between generations in this way (i.e., specifying stimulus IOIs by the previous participant's tempo instead of their ITIs) served to isolate the tempo of reproductions.

Procedure. Each participant performed four types of tapping tasks, all of which required tapping at a steady rate on the response pad. The four types of tapping tasks fell into two categories (described further below), which were *self-paced tapping* and *serial reproduction (SR)* tasks.

Self-paced tapping tasks. Each participant generated several sequences of self-paced taps at a comfortable tempo, a fast tempo, and a slow tempo. The three self-paced tapping tasks assessed spontaneous motor preferences and the tempo limits of rhythm production ability. In the spontaneous motor tempo (SMT) task, participants tapped with the index finger of their

34

dominant hand at a pace that was neither too fast nor too slow but that felt just right or comfortable to them. Participants tapped thirty-one times at this comfortable rate, and performed the task three times in three separate blocks distributed throughout the experiment session.

The two other self-paced tapping tasks were intended to assess the tempo limits of participants' rhythm production abilities. In the fast tempo limit task, participants tapped with their finger thirty-one times as quickly as possible while maintaining a steady beat. In the slow tempo limit task, participants tapped thirty-one times as slowly as possible while maintaining a smooth and continuous rhythm. Participants performed two trials of each of the fast and slow tempo limits tasks, distributed throughout the experiment session.

Serial reproduction (SR) tapping task. The SR task consisted of an exposure phase, in which participants were presented with the isochronous stimulus rhythm, and a reproduction phase, in which participants reproduced the tempo of the rhythm by tapping. Participants performed an exposure-reproduction trial for each chain three times (in three separate blocks). Within each block, the order of presentation of the chains was random.

The exposure task involved either aligning finger taps with the onset of events in the stimulus rhythm (synchronize condition), or passively observing the stimulus rhythm (observe condition). During either synchronize or observe conditions of the exposure phase, participants in the visual condition were presented with the visual stimulus rhythm (the sequence of green circles) on the screen of the experiment computer. Additional text at the bottom of the screen said either *tap along* (in the synchronize condition) or *observe* (in the observe condition). During either synchronize or observe conditions of the exposure phase in the auditory condition, participants heard the auditory stimulus rhythm (the sequence of tones), and simultaneously saw either a static *GO* sign with text that said *tap along* (in the synchronize condition) or a static

*LISTEN…*sign (in the observe condition). In all exposure conditions, after thirty-one presentations of the stimulus event (circle or tone), the sequence ended and participants were instructed to tap at the same rate as the sequence (reproduction phase). In all conditions, during the reproduction phase when participants reproduced the timing of the stimulus by tapping, a static *GO* sign identical to the one used during the synchronize condition was displayed on the screen. The reproduction phase of the experiment was identical across all conditions.

Experimental session structure. For all participants, the experiment session was split into three blocks, separated by two breaks. During breaks, the computer did not allow the participant to continue for a minimum of one minute, after which the participant pressed the spacebar to proceed with the experiment when ready. The first block consisted of all tasks, in the following order: SMT, fast tempo limit, slow tempo limit, and SR. The second block consisted of only the SMT and SR tasks. The third and final block consisted of SMT, SR, fast tempo limit, and slow tempo limit tasks. The order of tasks aimed to minimize potential carryover effects from the stimulus rates presented during the SR task to the SMT, fast tempo limit, or slow tempo limit tasks.

After the experimental (tapping) trials, participants completed a brief questionnaire about demographics and strategies. The entire session lasted approximately one hour.

Stimulus construction between participants. After each participant session, stimuli for the next participant in the cohort were constructed based on the previous participant's reproductions. Only responses from the reproduction phase were iterated across participants. Each participant performed three reproductions for each of seven chains. The median of the ITIs from each reproduction for each chain was calculated, and then the mean of individual medians

36

was calculated. The means of median ITIs (i.e., the tempi) from one participant's reproductions were used as the IOIs for the isochronous stimulus rhythms presented to the next participant[1].

Data Analysis

Self-paced tapping tasks. There were three trials of SMT and two each of fast and slow tempo limits tasks. Data were first pre-processed, and then averaged to calculate values for SMT, fast, and slow tempo limits for each participant, as detailed in the next two sections.

Preprocessing of self-paced tapping data. A trial was excluded if it included fewer than 10 (out of the target 30) ITIs. Eleven trials with fewer than 10 ITIs were excluded based on this criterion (leading to exclusion of 18 out of 64215 ITIs). An individual ITI was excluded if it was either extremely fast (defined as < 100 ms, a potential accidental double-tap) or extremely slow (defined as > 1.75*block median, a potential accidental miss by the participant or tapping device). Less than 0.5% of all ITIs (165/64215) were removed for being below the lower-bound of 100 ms. The upper-bound was based on an individual participant's trial median (rather than a constant value, as was done for the lower-bound cut-off) in order to catch equipment error (missed taps) but allow for genuine slow-tapping by some individuals on the slow tempo limits task (e.g., a 3000 ms ITI might be "noise" on a trial otherwise consisting of mostly ~500 ms ITIs, but be "signal" on a trial otherwise consisting of other ~3000 ms ITIs). The particular criterion of 1.75*median was intended to cut out obvious missed taps (e.g., an ITI twice as long as the

---

[1] Prior to averaging across reproductions, medians from each reproductions for each chain were qualitatively compared to each other, and to the target stimulus tempo, in order to identify outlier reproductions. For example, if the medians from the first, second and third reproductions were 925, 950, and 250 ms, respectively, in responding to a 1000 ms stimulus, then I flagged the third reproduction as a potential outlier and further inspected the time-series of ITIs for each reproduction for that participants. Inspecting the time-series of flagged reproductions in most cases revealed equipment or participant error for that trial. Equipment error often demonstrated a characteristic pattern, including many ITIs < 100 ms (indicating electrical noise in the equipment). Another common pattern in the time-series that I took to indicate missed taps by the device (either due to the participant tapping too lightly for the device to detect, or not tapping) was one or more multiple-second ITIs (which, in turn, led the reproduction to have very few ITIs contributing to the median calculation). When the reproduction's outlier status was clearly attributable to this type of error based on visual inspection of the time-series, I excluded the trial from the mean tempo calculation

average ITI would seem to reflect maintenance of the target tempo, but an accidental miss), but also catch slightly 'early' missed taps, while still being a fairly liberal inclusion criterion. Approximately 3.4% of all ITIs (2162/64215) were removed based on the upper-bound cut-off of 1.75*block median ITI.

Spontaneous motor tempo (SMT). Similar to Experiment 1, the median of the ITIs from each SMT production were calculated for each participant. In Experiment 2, there were three SMT productions. The mean of the individual production medians was then calculated and served as the SMT for that participant (following the method of McAuley et al., 2006). Calculating the mean of multiple trial medians was intended to provide a more reliable estimate of SMT relative to using a single trial median.

In order to establish the expected preferred tempo of the experiment sample, the mean SMT across all participants in all conditions was calculated, averaging over participant SMTs. As chains were not expected to converge in all conditions, the average SMTs across participants by condition were also calculated, in order to assess the expected preferred tempo by condition.

Fast and slow tempo limits. There were two productions each of the fast and slow tempo limits tasks. Similar to the method for SMT, the medians of ITIs from each production were calculated and then averaged in order to provide the fast and slow tempo limits for each participant. In addition, the fast and slow tempo limits were used to calculate normalized produced range (NPR), an indicator of the width of an individual's entrainment region (McAuley et al., 2006). NPR was calculated by finding the difference between the slow and fast tempo limits, divided by the participant's SMT (i.e., NPR = (slow tempo limit – fast tempo limit) /

(in which case the mean tempo calculation was based on the remaining reproductions). Twenty-nine trials (out of 6300 total) were excluded by this method.

SMT). The mean fast and slow tempo limits and NPR were calculated by cohort, averaging across participants, in order to compare to performance on the SR task.

Serial reproduction (SR) chains. Each participant made three reproductions for each of seven chains. The medians of the ITIs from each reproduction were calculated, and then the mean of the medians was calculated to determine the tempo to use for the next participant's stimulus rhythms.

Drift of serial reproductions. Drift in SR chains was calculated in a similar manner to Experiment 1 by finding the difference between the final tempo in a chain and the seed tempo of the chain. Linear regression was used to find the equation describing the change in drift as a function of seed tempo by condition. A negative slope indicates that chains converged. In this case, the x-intercept estimates the convergence tempo of the chains in the condition. A zero slope indicates that chains did not converge. One-sample t-tests were conducted to test whether slopes differed significantly from zero.

The slope of the linear equation was of primary interest in evaluating the effects of the input noise manipulations in this experiment. Conditions with lower levels of input noise (and thus less drift) were predicted to yield slopes closer to zero than conditions with higher levels of input noise (and thus more drift). The 99% confidence intervals (CIs) for slope values in each condition were constructed by calculating a margin of error for each slope estimate (CI = slope $\pm$ margin of error). The margin of error was calculated by finding the critical value (based on a t-score with n-2 degrees of freedom, where $n = 35$, the number of points in each regression, for all conditions in the experiment) and multiplying by the standard error of the slope. The lower and upper CIs are the slope plus or minus (respectively) this margin of error (Cohen, Cohen, West, & Aiken, 2013). The overlap of a slope estimate with the CI of another slope estimate suggests that

the two slopes do not differ reliably. However, in order to further test the difference in slopes across conditions, several independent t-tests were conducted (Cohen et al., 2013; Paternoster, Brame, Mazerolle, & Piquero, 1998).

Within each exposure condition, independent sample t-tests were conducted on regression slopes in order to compare whether the effect of observing or synchronizing with the stimulus during exposure differed across stimulus modalities. Specifically, the first comparison tested the prediction that slopes in the visual-observe condition would be steeper (more negative) than slopes in the auditory-observe condition. The second comparison tested the prediction that the slope in the visual-synchronize condition would be more negative than the slope in the auditory-synchronize condition (though both slopes were generally predicted to be non-significantly different from zero).

Relationship between SR and SMT. In order to evaluate whether chain convergence tempo reflected the preferred tempo of individual cohorts, a new measure, relative tempo, was calculated. Relative tempo is the difference between stimulus tempo and SMT divided by SMT (i.e., (stimulus tempo – SMT) / SMT). Similar to NPR (described above), relative tempo is a re-scaling of stimulus tempo in "SMT units". One "unit" of relative tempo equals one SMT. A relative tempo equal to zero correspond to a stimulus tempo equivalent to the participant's SMT. Stimulus tempi faster than SMT have negative relative tempo, and conversely stimulus tempi slower than SMT have positive relative tempo.

Linear equations were fit to drift as a function of relative seed tempo by condition, and also by cohort. Fitting by condition allowed for comparison to the parallel analysis using seed tempo in millisecond units. Fitting by cohort enabled assessment of the consistency of the by-

condition fits across cohorts, as well as assessment of whether convergence tempi matched preferred tempi on a more fine-grained level.

Results

Spontaneous motor tempo (SMT) and tempo limits. Table 2 summarizes the fast tempo limit, SMT, slow tempo limit, and normalized produced range (NPR) data across participants for each of the four experimental conditions. To assess the preferred tempo of the entire sample, the mean SMT across all conditions was calculated ($N = 300$). Mean SMT was 630.1 ms ($SD = 199.1$), which is consistent with previous literature (Drake et al. 2000; McAuley et al., 2006). The mean fast tempo limit ($M = 196.2$ ms, SD = 44.2) and slow tempo limit ($M = 1760.6$ ms, $SD = 855.0$) were also fairly consistent with previous research.

Table 2. Experiment 2: Descriptive summary of SMT and tempo limits by condition

		Auditory		Visual	
		Observe	*Synchronize*	*Observe*	*Synchronize*
Fast	*M*	192.4	184.8	206.7	200.8
	SD	46.3	27.2	49.2	47.9
SMT	*M*	617.0	655.4	613.9	634.2
	SD	173.6	218.1	211.4	191.5
Slow	*M*	1630.5	1987.9	1688.8	1734.3
	SD	945.8	860.0	845.5	727.6
NPR	*M*	2.4	3.0	2.5	2.6
	SD	1.8	1.6	1.1	1.8

Note. Fast = fast tempo limit. SMT = spontaneous motor tempo. Slow = slow tempo limit. NPR = (Slow – Fast) / SMT. M = mean. SD = standard deviation. For Fast, SMT, and Slow, units are milliseconds.

Serial reproduction. Analysis of the SR chains paralleled Experiment 1. In Figure 5A, the grey lines show individual chains of tempo reproductions. Seed tempo corresponds to

iteration zero, and iteration (serial position in the chain) is shown on the x-axis. Colored lines show the mean tempo averaged across individual chains initiated at the same seed tempo. In the observe condition, regardless of modality, chains initiated with different seed tempi appeared to drift towards a common tempo, whereas chains in the synchronize condition demonstrated reduced drift, maintaining more of the tempo spread present initially in the seed stimuli.
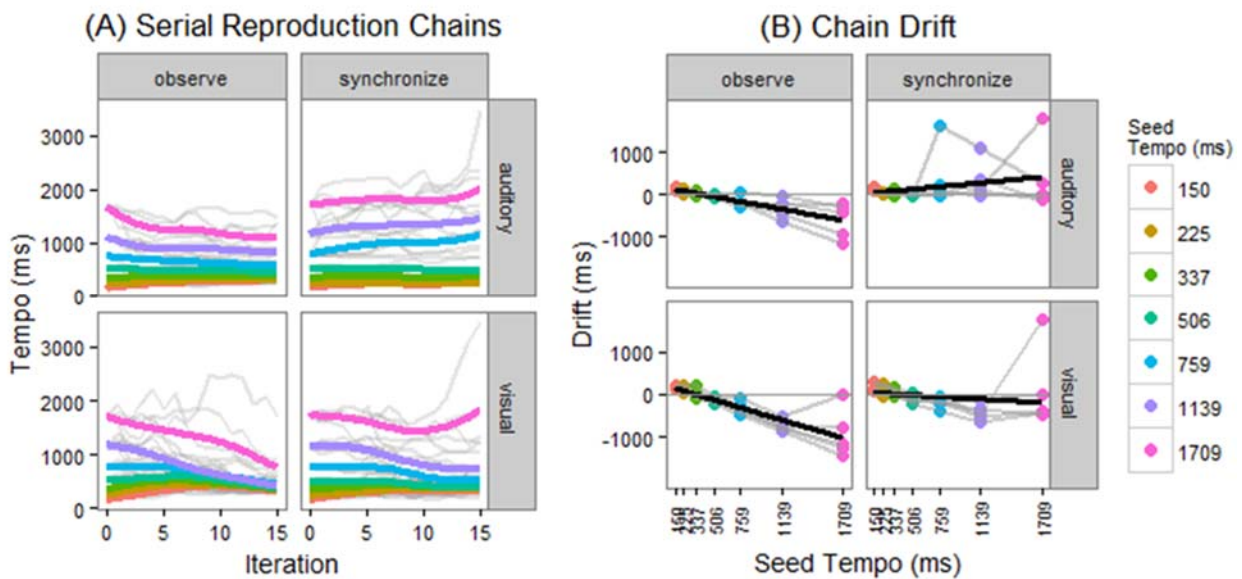
Figure 5. Serial reproduction by condition in Experiment 2
Panel A shows tempo reproductions for each modality (auditory, visual) by exposure (observe, synchronize) condition. Light grey lines plot data from individual chains, with colored lines summarizing the mean tempo of reproductions by different participants at each iteration for chains initiated at the same seed tempo. In Panel B, chain drift from seed to final tempo is plotted as a function of seed tempo. Each point represents data from an individual chain (color codes for the seed that initiated the chain summarized by the point), with chains belonging to the same cohort connected by grey lines. The solid black lines represent the best-fit line to drift as a function of seed tempo across cohorts in a condition. For lines with negative slope, the intersection of the best-fit line with the x-axis corresponds to the predicted convergence tempo of the chains in that condition.

As in Experiment 1, the change in tempo from seed stimulus to final reproduction for each chain was formalized as drift, which is the signed difference between the final and seed

tempi. As shown in Figure 5B, drift was calculated for each SR chain and then regression lines for each experimental condition were fit to drift as a function of seed tempo. The results of the regression analysis are summarized in Table 3 (left columns under the heading "Drift ~ Seed Tempo"). Consistent with the predictions of the preferred tempo and input noise hypotheses, slopes were negative and significantly different from zero in both auditory- and visual-observe conditions, whereas slopes were not significantly different from zero in either auditory- or visual-synchronize conditions. For the conditions in which the slope was significantly non-zero, the x-intercept of the regression line was calculated to ascertain the convergence tempi of chains in that condition. For the auditory-observe condition, the estimated convergence tempo was 411.0 ms. For the visual-observe condition, the estimated convergence tempo was 352.9 ms.

To further evaluate the predictions of the input noise hypothesis, two additional independent-samples t-tests comparing the effect of stimulus modality within each exposure condition were conducted. Within the observe condition, a t-test comparing slopes for auditory and visual conditions revealed that chains converged more quickly in the visual condition, as predicted by the input noise hypothesis, $t(66) = 3.07$, $p = .003$. Within the synchronize condition, a t-test comparing slopes for auditory and visual conditions revealed that chains converged more in the visual condition (based on a negative slope that was, however, not significantly different from zero) compared to chains in the auditory condition (in which the slope was both positive and not significantly different from zero), $t(66) = 2.05$, $p = .04$.

Relationship between SR and SMT. The analysis so far sought to establish that SR chains converged to a common tempo, hypothesized to reflect average preferred tempo, regardless of seed tempo, and further that convergence rates depended on input noise for the stimulus. The next analysis investigated whether chain drift reflected, specifically, the average

preferred tempo of the participants in the chain by calculating the mean relative tempo by cohort. Relative tempo is the difference between stimulus tempo and the average SMT of the cohort, divided by the cohort SMT. One unit of relative tempo is equivalent to the value of the mean SMT for a cohort. A relative tempo equal to zero indicates that the stimulus tempo is equivalent to SMT.

Table 3. Experiment 2: Drift equations by condition

	Auditory		Visual	
	Observe	*Synchronize*	*Observe*	*Synchronize*
Drift ~ Seed Tempo				
Slope	-0.47***	0.24	-0.75***	-0.15
CI	-0.63, -0.31	-0.13, 0.61	-0.95, -0.55	-0.50, 0.21
Y	192.2	4.6	265.6	63.3
X	411.0	-19.2	352.9	433.6
R²	0.67	0.09	0.77	0.04
Drift ~ Relative Seed Tempo				
Slope	-302.9***	198.0*	-476.6***	-60.1
Y	-93.2	157.2	-190.0	-31.8
X	-0.31	-0.79	-0.40	-0.53
R²	0.75	0.15	0.85	0.02

Note. CI = lower and upper bounds of the 99% confidence interval for the slope. Y = Y-intercept from linear fit. X = X-intercept calculated from linear fit. The equation for the "Seed Tempo" fit in the left columns was *Drift (ms) = Slope * Seed Tempo (ms) + Y*. The equation for the "Relative Seed Tempo" fit shown in the right columns was *Drift (ms) = Slope * Relative Seed Tempo + Y*. Asterisks indicate significant difference from zero * $p <$ .05, ** $p <$ .01, *** $p <$ .001.

Figure 6A shows the change in relative tempo across reproductions. In order of increasing value, the dashed horizontal lines indicate the fast tempo limit, SMT, and slow tempo limit (in SMT units) by condition. Drift towards SMT on this plot is reflected by convergence of chains towards zero. As predicted by the preferred tempo hypothesis, chains in the observe condition appeared to drift to zero in both auditory and visual conditions. Consistent with the

input noise hypothesis, chains in the synchronize condition did not appear to converge in the auditory condition and appeared to converge more slowly in the visual condition.

Figure 6B shows the relationship between drift and relative tempo. Paralleling the analysis with millisecond stimulus tempi, regression lines for each experimental condition were fit to drift as a function of relative seed tempo. The results of the regression analysis are summarized in Table 3 (right columns under the heading "Drift ~ Relative Seed Tempo"). As expected, in the observe condition in both modalities, drift magnitude varied significantly as a negative linear function of relative seed tempo. Consistent with the preferred tempo hypothesis, the estimated convergence tempi of chains in both the auditory and visual observe conditions were close to preferred tempo as measured by SMT, which, in relative tempo units, was zero. Consistent with the input noise hypothesis, drift towards preferred tempo was reduced in the synchronize condition in both modalities. The slope in the auditory-synchronize condition was significantly different from zero, but positive, suggesting that chains in this condition diverged, rather than converging. Consistent with the predicted interaction of modality and exposure, the slope in the visual-synchronize condition was slightly negative, though not significantly different from zero.

Figure 7 plots estimates of slope and x-intercept for individual cohorts by the filled colored dots. For comparison, the fits by condition are plotted by the empty circles. For the most part, the convergence rates (slopes) and tempi (x-intercepts) obtained for the condition fits were consistent across cohorts within a condition. All cohorts in both auditory and visual observe conditions had negative slopes and convergence tempi near zero, consistent with the preferred tempo hypothesis and the condition-level fit. For the auditory-synchronize condition, three cohorts had slopes near zero, as predicted by the input noise hypothesis and consistent with the

condition-level fit. However, two cohorts in this condition had positive slopes, indicating that

chains initiated at different seed tempi diverged for these cohorts. On average, these two cohorts

also had larger NPR. The finding is consistent with the possibility having a wider entrainment

region led chains in these cohorts to diverge, although it is too little data to say conclusively. For

the visual-synchronize condition, most cohorts had negative slope (indicating convergence) and a

convergence tempo around zero, with one cohort instead having a positive slope (indicating

divergence).
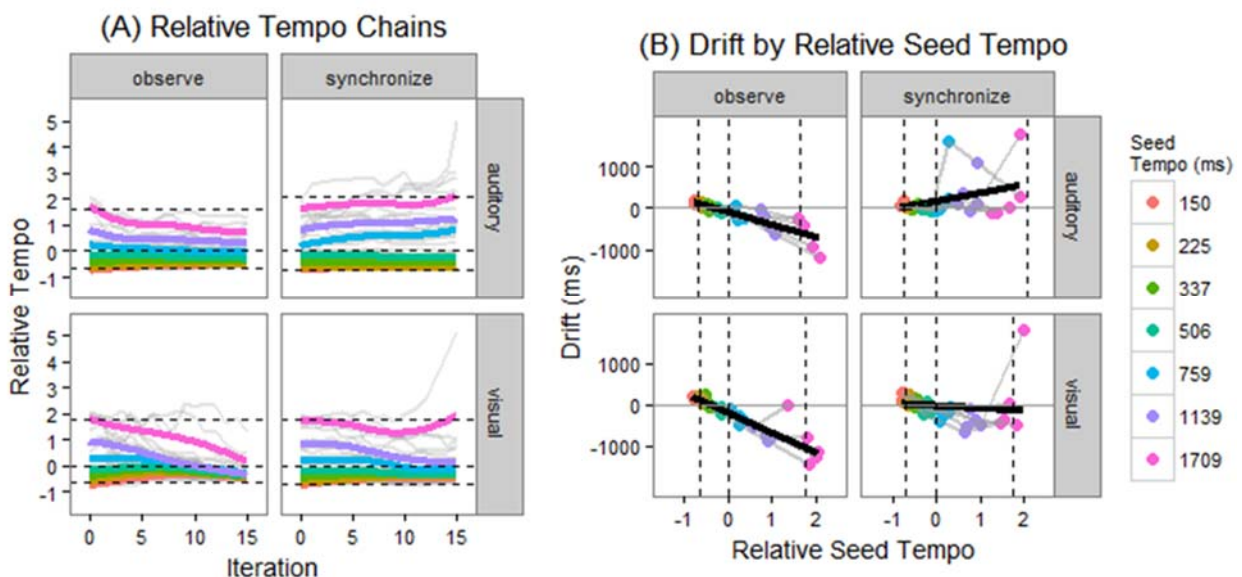
Figure 6. Relative tempo chains and drift in Experiment 2
Panel A shows tempo reproductions in relative tempo units. Relative tempo was found by calculating (Stimulus Tempo – SMT) / SMT for each cohort. Grey lines show data for individual chains; colored lines plot the mean relative tempo, summarizing over cohorts, for chains initiated by the same seed tempo. The dashed horizontal lines correspond to mean self-paced tapping tempi in SMT units calculated for each condition. The line at $y = 0$ denotes the stimulus tempo that is equivalent to SMT. The horizontal lines above and below this denote the mean fast and slow (respectively) tempo limits by condition. Panel B show chain drift as a function of relative seed tempo. From left to right, the dashed vertical lines correspond to fast tempo limit, SMT, and slow tempo limit (in SMT units), calculated by condition. The black lines show the best fit lines through the data by condition. For conditions with negative drift slopes, the estimated convergence tempo of chains in each condition is the intersection of this line with the x-axis.
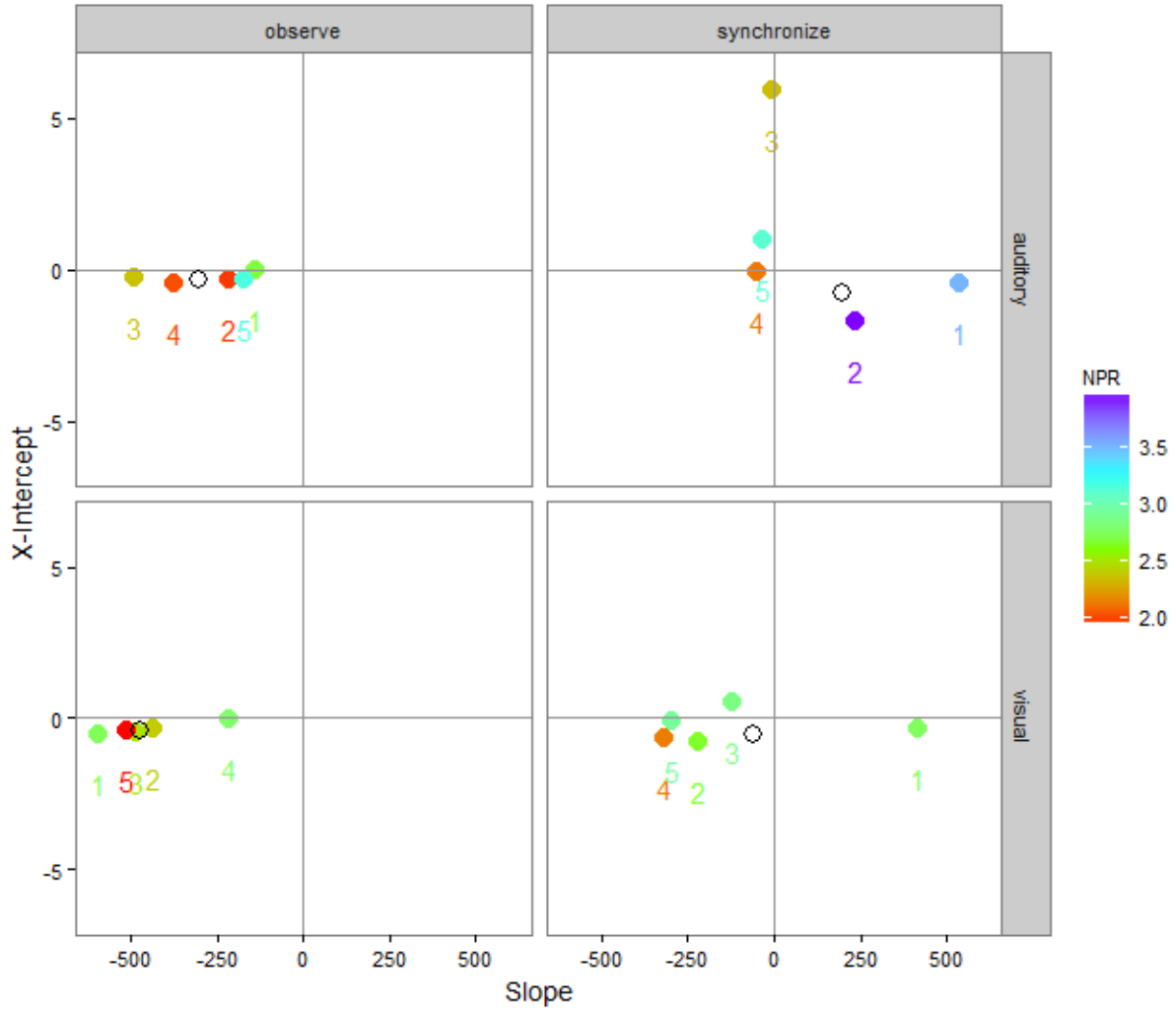
Figure 7. Convergence rates and tempi in Experiment 2
The estimates of convergence rate (slope) and tempo (x-intercept) based on fitting a regression line to drift as a function of relative tempo are plotted for each cohort. Number labels indicate the identifier for the cohort. Empty circles correspond to fits by condition. Slopes less than zero indicate that chains initiated by different seed tempi converged to the x-intercept. Slopes equal to zero indicate that chains did not converge. Slopes greater than zero suggest that chains diverged. An x-intercept equal to zero indicates that chains converged to a tempo equivalent to SMT. X-intercepts below or above zero indicate convergence to a tempo faster or slower, respectively, than SMT. The color of the dot corresponds to the average normalized produced range (NPR) of the cohort. NPR = (slow tempo limit – fast tempo limit) / SMT. A larger value of NPR indicates a wider range of entrainable tempi.

Discussion

Experiment 2 had two main goals, which were (1) to test the preferred tempo hypothesis and (2) to test the input noise hypothesis. The preferred tempo hypothesis proposed that an inherent tempo preference based on properties of an oscillatory internal timekeeper influences the perceptual representation of rhythms, such that the perceived tempo of to-be-learned rhythms would be systematically distorted (i.e., biased) to be more similar to a preferred tempo. The input noise hypothesis proposed that uncertainty (noise) during initial encoding of a stimulus affects the amount of bias in the perceptual representation (and consequently, the reproduction) of the stimulus. Specifically, conditions resulting in greater input noise will lead to greater drift in SR chains than conditions with less noise. To test the preferred tempo hypothesis, Experiment 2 replicated and extended Experiment 1 by presenting participants in SR chains with a wider range of seed tempi in a laboratory setting. To test the input noise hypothesis, Experiment 2 varied the modality (visual, auditory) of stimulus rhythms and exposure task (synchronize, observe) of participants during exposure to the stimulus rhythms.

As predicted by the preferred tempo hypothesis, in both modalities in the observe condition, fitting a line to chain drift as a function of seed tempo revealed negative and significantly non-zero slopes, consistent with the prediction that chains in these conditions would converge. The estimated convergence tempo for auditory chains was 411 ms and for visual chains was 353 ms. Convergence tempi were, thus, in the predicted direction (i.e., fast-seed chains slowed down and slow-seed chains sped up), although somewhat faster than expected based on the average SMTs in the two conditions, which were 617 ms and 614 ms, respectively. On average, chains initiated by the three fastest seeds (150, 225, and 337) tended to slow down, whereas chains initiated by the three slowest seeds (759, 1139, and 1709) tended to speed up.

48

Chains initiated by the intermediate seed (506) changed the least. This pattern of tempo change across reproductions is consistent with a description of tempo bias as the attraction of tempo representations towards the equilibrium state (i.e., the inherent period) of an oscillatory internal timekeeper that is around 500 to 600 ms.

The finding that the estimated convergence tempo was slightly faster on average for the visual compared to the auditory chains suggests that there may be a bias towards a faster preferred tempo for visual compared to auditory rhythms. This result is somewhat surprising given the previous literature comparing discrimination and production performance in response to auditory and visual rhythms, which generally suggests that the preferred processing tempo for visual rhythms should be slower than for auditory rhythms (Collier & Logan, 2000; McAuley & Henry, 2010; Repp, 2003). At this time, it is not entirely clear how to explain this result, which merits further investigation.

As a further test of whether chains converged specifically to the preferred tempo of the participants tested, a relative tempo measure was introduced, which converted stimulus tempi into "SMT units". Relative stimulus tempo was calculated by finding the difference between the stimulus tempo and SMT, and then dividing by SMT. In units of relative tempo, convergence to zero indicates that reproductions were biased towards SMT. Thus, chains in the observe condition were predicted to converge to zero. Consistent with this prediction, a linear fit to drift as a function of relative seed tempo revealed that chains in both auditory and visual observe conditions converged, on average, to a relative tempo close to zero. Fitting drift as a function of relative seed tempo for each cohort revealed that all cohorts in auditory and visual observe conditions converged to a relative tempo of close to zero, as well, consistent with predictions.

Consistent with the input noise hypothesis, chain drift towards preferred tempo was reduced in the synchronize compared to the observe condition, for both auditory and visual rhythms. Drift slopes in the synchronize condition were not significantly different from zero, suggesting that synchronizing with the rhythm during exposure reduced initial encoding noise for the stimulus, shifting perceptual representation towards the stimulus tempo. In contrast, drift slopes in the observe condition were negative and significantly different from zero, suggesting that higher initial encoding noise for the stimulus in this condition led perceptual representations to rely to a greater degree on preferred tempo.

Based on previous work showing that moving in synchrony with a rhythmic stimulus can improve the accuracy of timing judgments (Arnal et al., 2014; Iordanescu et al., 2013; Manning & Schutz, 2013), pitch judgments (Morillon et al., 2014) and pulse extraction (Su & Pöppel, 2012), as well as enhancing neural response to periodic stimuli (Nozaradan et al., 2016; Tierney & Kraus, 2013), it was expected that synchronizing finger taps with the stimulus during exposure would reduce input noise relative to passively observing the stimulus. Finding flatter slopes in the synchronize vs. observe condition supports the proposal that movement benefits the accurate representation of temporal information by reducing initial encoding noise (Morillon et al., 2016; Nozaradan et al., 2016; Schroeder et al., 2010). The reduction in bias in the synchronize condition suggests that movement not only increased the accuracy of temporal sensory representations (consistent with past proposals), but further, served to reduce bias in perceptual representations (i.e., beyond 'sharpening' sensory representations, movement potentially shielded perceptual representations from systematic bias). This result is consistent with the possibility that increasing the reliability of early input representations shifts implicit weighting during perceptual representation from preferred tempo to recent stimulus memory, because

moving in synchrony with the stimulus rhythm facilitates veridical initial encoding of that rhythm (Morillon et al., 2016; Nozaradan et al., 2016; Schroeder et al., 2010).

Based on the previous findings of general performance advantages for perceiving and reproducing auditory relative to visual rhythms (Collier & Logan, 2000; Grondin & McAuley, 2009; Patel et al., 2005; Repp, 2003, 2005; Repp & Su, 2013), presentation of stimulus rhythms in the visual modality in the present experiment was expected to increase input noise relative to the auditory modality. Increased input noise of visual relative to auditory rhythms was predicted to lead visual SR chains to converge more quickly than auditory SR chains. Consistent with this prediction, visual chains in the observe condition converged more quickly (as evidenced by a larger-magnitude slope) than auditory chains in the observe condition, suggesting that input noise may have been greater for the visual than the auditory rhythms.

Drift slopes in the synchronize condition were not significantly different from zero in either modality. However, the pattern of drift in individual cohorts tended to be zero or slightly positive in the auditory-synchronize condition (consistent with reduced input noise), whereas the pattern of drift in individual cohorts tended to be negative in the visual-synchronize condition, suggesting that synchronization with visual rhythms may not have reduced input noise as consistently as it did for auditory rhythms. Together, the difference in the magnitude of convergence rates along with differences in the effect of synchronous movement on convergence rates across modalities contributes to increasing evidence of noisier initial representations for visual compared to auditory rhythmic information (Becker & Rasmussen, 2007; McAuley & Henry, 2010; Recanzone, 2003; Repp, 2003; Repp & Penel, 2002, 2004).

The preferred tempo hypothesis fits within a broader entrainment-based perspective on timing and temporal processing. From an entrainment perspective, preferred tempo corresponds

to the inherent oscillation period of an internal timekeeper. The period of the oscillator transiently adjusts to the period of repetitive external stimulation, but drifts back to its inherent period (i.e. equilibrium state) in the absence of external input. Each SR chain is akin to mapping the return of the theoretical internal oscillator period from a perturbed state (the tempo of the seed rhythm) to its equilibrium state (preferred tempo) by taking snapshots of its state (the tempo reproductions) at discrete time steps (each iteration). Thus, SR chains of tempo were predicted to drift towards preferred tempo, regardless of seed tempo, because preferred tempo is a stable, equilibrium state of the system.

One way to understand the effect of input noise on tempo bias within an entrainment framework is to interpret the noise associated with a stimulus tempo as influencing the amount of period adjustment made by the internal oscillatory timekeeper in response to external stimulation (McAuley, 1995; McAuley & Jones, 2003). In an entrainment framework, internal periodicities adjust in response to external periodic stimulation, and in that way sequence tempo is represented by the adjusted period of the internal oscillation. However, adjustment of the internal periodicity in response to the stimulus tempo may be incomplete – that is, the period may come to fully match the stimulus tempo (full adjustment), or may partially adjust, thereby exhibiting bias based on its inherent period and based on foregoing temporal context (McAuley & Jones, 2003).

The present results potentially identify one factor (i.e., input noise) that influences the amount of adjustment towards stimulus tempi made by an internal oscillatory timekeeper. Greater period adjustment to stimulus tempo will lead representations to reflect the tempo of the stimulus, whereas decreased period adjustment will lead representations to reflect (to a greater degree) the inherent period of the oscillator (i.e., preferred tempo). Considered within the SR

paradigm, this predicts that chains of reproductions in which input noise associated with the stimulus is relatively low will reflect the tempo of the stimuli (due to greater adjustment of internal periodicities), whereas chains of reproductions in which input noise is relatively high will reflect drift towards the tempo of the internal oscillator (due to decreased adjustment of internal periodicities). The lack of systematic drift by chains in the synchronize conditions, but significant drift by chains in the observe condition, is consistent with this characterization.

An alternative explanation for the lack of drift in at least the synchronize condition comes from interval models of timing (Church, Miller, Meck, & Gibbon, 1991; Gibbon, 1977; Gibbon, Church, & Meck, 1984; Treisman, 1963). From an interval point of view, tempo is represented as the running average of the perceived duration of stimulus IOIs (Drake & Botte, 1993), each of which is centered on the true value of the stimulus, without bias (Wing & Kristofferson, 1973a, 1973b). In general, interval-based models predict that the inter-tap intervals of the reproduction of an isochronous stimulus will directly reflect the inter-onset intervals of the stimulus, without bias, because the means of both the stored code for duration and the reproduced interval are centered on the stimulus IOI. With regards to the present experiment, an interval-based model such as that of Wing & Kristofferson (1973a, 1973b) does not predict any systematic error in the reproduced interval relative to the stimulus interval, which is consistent with the flat slopes found in the synchronize condition. The results in the observe condition, in contrast, seem generally inconsistent with the interval approach, since SR chains were systematically biased towards preferred tempo.

One question raised by these results is why the predicted convergence tempi for auditory and visual SR chains in the observe condition (~350 to 400 ms) were faster than the tempo expected based on the average SMT in these conditions (~617 ms). One possible origin for this

53

discrepancy could be the various strategies participants used to improve tracking of stimulus tempo during the experiment. Of the participants who reported using a strategy to help remember stimulus tempi, several participants reported grouping or subdividing events in the stimulus rhythm. If this strategy caused group-length, rather than ITI, to be biased towards preferred tempo, that would lead chain convergence tempi to be faster than expected. For example, if the group-length of a pair of ITIs was biased towards 630 ms, then the reproduced ITI would be biased towards 315 ms.

Another reason that chains converged to a faster than expected tempo could be that averaging over participants distorted the estimate of chain convergence tempo. One issue with basing the expected convergence tempo of chains on the average SMT of the sample is that it assumes that preferred tempo is the same across participants, which is only accurate on average. Individual differences in preferred tempo and entrainment region have been observed previously (Drake et al., 2000; McAuley et al 2006). Having separate individuals contribute reproductions at each iteration of an SR chain allows for the possibility that perturbations (such as those based on differences in SMT or in strategy) may have been introduced.

In order to control for variation in individual differences in preferred tempo, strategy use, or other potential factors, Experiment 3 replicated the auditory condition of Experiment 2 using single participant SR chains. As in Experiment 2, the preferred tempo hypothesis predicted that single-participant chains would converge to preferred tempo, as measured by SMT. The input noise hypothesis predicted that convergence rates would be reduced when participants synchronized with the stimulus rhythms.

More broadly, Experiment 3 provided a further test of whether the SR paradigm provides a useful tool for examining the mental representation of rhythm. Generally, SR chains are

predicted to converge to the underlying expectation that biases reproductions at each iteration (in this case, preferred tempo), regardless of whether the chains consist of reproductions by multiple participants or by a single participant. The utility of the SR method (in contrast to other methods for examining individual differences) is that small biases may be amplified by serial reproduction.

Conclusion

In Experiment 2, serial reproduction chains of both auditory and visual rhythms initiated at seed tempi spanning 150 to 1709 ms converged to a tempo of about 350 to 400 ms when participants passively observed the stimulus rhythms. In contrast, when participants synchronized with the stimulus rhythms, drift in both auditory and visual serial reproduction chains was reduced, although the reduction in drift was greater for auditory than visual rhythms. Overall, these results support both the preferred tempo and the input noise hypotheses. One question raised by the present experiment was whether a similar pattern of results would be obtained using serial reproduction chains in which all reproductions were made by a single participant. Thus, in Experiment 3, participants either passively observed or synchronized with auditory stimulus rhythms in single-participant SR chains. Finding results consistent with the preferred tempo and input noise hypotheses in Experiment 3 will suggest that the method is a viable tool, not just for examining tempo bias at a group level, but for examining individual differences in tempo bias.

CHAPTER 4: EXPERIMENT 3. THE EFECT OF SENSORY ENCODING NOISE ON TEMPO BIAS IN SINGLE-PARTICIPANT SERIAL REPRODUCTION CHAINS

Experiment 3 tested the preferred tempo and input noise hypotheses using single-participant serial reproduction chains. Using single-participant chains enabled a more direct investigation of the role of individual differences in tempo biases. Moreover, this experiment assessed whether the pattern of results obtained using multiple-participant chains in Experiments 1 and 2 would generalize to the case where the same participant provided rhythm reproductions at each iteration of the chain.

As in Experiment 2, input noise was varied by manipulating exposure task (*observe* or *synchronize*) between groups of participants. Each participant was initially presented with auditory isochronous stimulus rhythms at the same seed tempi as Experiment 2, which they then reproduced by tapping. However, after the initial stimulus, participants were presented with isochronous stimuli at tempi based on their own immediately-preceding reproduction (rather than the reproduction of another participant) using a similar method to Experiment 2. Also, as in Experiment 2, participants additionally produced multiple self-paced tapping sequences at comfortable, fast, and slow steady tempi in order to measure participants' spontaneous motor tempo (SMT), fast tempo limit, and slow tempo limit.

Based on the preferred tempo and input noise hypotheses, the pattern of chain drift in Experiment 3 was predicted to parallel the pattern found in Experiment 2. That is, SR chains were predicted to converge to preferred tempo regardless of seed tempo, with convergence reduced in the synchronize compared to the observe condition. With regards to the preferred tempo hypothesis, SR chains in Experiment 3 were predicted to converge to the preferred tempo of the single participant whose reproductions constituted each reproduction in the chain (as

measured by the participant's SMT). This was expected to be especially true in conditions with greater input noise, in which successive reproductions were expected to rely to a greater degree on inherent tempo biases.

One question that emerged from Experiment 2 was why the mean convergence tempo was somewhat faster than expected based on the mean SMT of participants in the experiment. This was conjectured to potentially have resulted from individual variability in the preferred tempo of participants, or potentially from differences in grouping strategy across participants – for example, some participants reported imagining the events in stimulus rhythms as grouped into pairs or trios, thereby potentially tapping at a faster rate if the group length was biased towards preferred tempo instead of the ITI. Thus, in replicating Experiment 2 using single-participant chains, Experiment 3 additionally aimed to clarify whether the faster chain convergence tempo (i.e., x-intercept of the drift function) was a genuine reflection of the mean preferred tempo of the participants, or a by-product of iterating reproductions across variable participants. A single participant was expected to be more consistent in both preferred tempo (Fraisse, 1982) and in strategy use than multiple participants.

As in Experiment 2, the relationship between the fast and slow tempo limits and tempo bias in the SR task was explored by calculating normalized produced range (NPR, McAuley et al., 2006). Results of Experiment 2 were consistent with the input noise hypothesis in that the majority of chains in the synchronize condition did not converge (slopes were not significantly different from zero). However, two cohorts had positive drift slopes, indicating divergence of chains. The atypical result for these two cohorts was hypothesized to potentially have resulted from these cohorts having, on average, wider NPRs than the other cohorts in the condition.

In Experiment 3, the consistency of chain convergence by single-participant chains with group-level estimates were evaluated with respect to participant NPRs. One outcome that was predicted was that participants with a wider NPR (indicating a wider band of tempi affording stable entrainment) would demonstrate weaker convergence than participants with a narrow NPR, based on the capability to accurately track a broader range of stimulus tempi.

Method

Design. The design was a 2 (Exposure task: synchronize, observe) x 7 (Seed tempo: 150, 225, 337, 506, 759, 1139, and 1709 ms) mixed-factorial design. Exposure was a between-subjects factor; seed tempo was a within-subject factor.

Participants. Twelve individuals ($N = 8$ female) from the Michigan State University community participated in the experiment for course credit or for pay at a rate of $10/hour. The mean age of participants was 27.3 years (range 18-45). For the participants who had musical training ($N = 6$), the number of years of training ranged from three to ten years ($M = 6.2$). All participants were right-handed native English speakers.

There was approximately the same number of chains per condition (i.e., thirty-five) in this experiment as in Experiment 2. Specifically, there were thirty-five chains in the observe condition (from five participants' reproductions in seven chains) and forty-nine chains in the synchronize condition (from seven participants' reproductions in seven chains). The difference in the number of participants between conditions resulted from arbitrary circumstances. The target minimum number of participants per condition was five, in order to match the number of chains in the parallel conditions in Experiment 2. However, due to scheduling and time constraints, more participants ended up both signing-up for and completing the experiment in the synchronize than the observe condition.

Equipment. Equipment was the same as Experiment 2.

Stimuli. As in the auditory condition of Experiment 2, stimulus sequences in Experiment 3 consisted of thirty-one repetitions of a 50 ms 400 Hz sine tone. As in Experiment 2, the stimuli for the first iterations of SR chains were sequences at the seed tempi (which were identical to those in Experiment 2). For the iterations that followed, the fixed IOI for the isochronous stimulus sequences in each chain were based on the median of the ITIs (i.e., the tempo) of the previous reproductions for each chain. In contrast to Experiment 2, the same participant provided reproductions at each iteration of a chain, with the median ITI from their own reproductions providing the tempo for the stimulus IOI in the next iteration of the chain.

Procedure. Participants in Experiment 3 performed the same self-paced and serial reproduction (SR) tapping tasks as participants in Experiment 2. As in Experiment 2, the SMT task involved tapping thirty-one times at a comfortable, steady rate. The fast and slow tempo limit tasks involved tapping thirty-one times as quickly or as slowly as possible, while still maintaining a sense of rhythm.

As before, the SR task consisted of an exposure phase, during which the participant synchronized with or passively observed (depending on condition) the stimulus rhythm, and a reproduction phase, during which the participant tapped to reproduce the timing of the stimulus rhythm. In contrast to Experiment 2, the tempi of the stimuli presented during the exposure phase were based on the tempo of the ITIs from the immediately preceding reproduction phase by the same participant. Between each chain iteration, stimuli were constructed by calculating the median of the ITIs from the preceding reproduction in each chain (with limited preprocessing[2]).

---

[2] The calculated tempo for an SR chain within an iteration was flagged as a potential outlier if there were too few recorded taps to calculate a median (there was one trial that fit this criterion), or if the calculated tempo was more than twice as slow as the stimulus tempo, indicating that there were many missed taps on that trial (two trials fit this

Participants were not told that the stimuli they were presented with came from their own reproductions.

Each participant performed fifteen blocks (one for each iteration in the SR chains), which each included all tasks in the following order: SMT, fast tempo limit, slow tempo limit, and SR. The fifteen blocks were evenly distributed across five sessions (three blocks per session). Blocks in the same session were identical to each other with the exception of minor modifications to instructions, which removed some redundancy. Sessions were nearly identical except for the questionnaire at the end. The first session concluded with a questionnaire about strategies and demographic information (including music and dance experience). Sessions two through five concluded with slightly briefer questionnaires about general state and any strategies used during that day. Each session lasted approximately one hour. Separate sessions were separated by a maximum of one week (and a minimum of one hour).

Data Analysis

Self-paced tapping tasks. In Experiment 3, each participant made fifteen productions for the SMT and tempo limits tasks (one production of each per block). The method for calculating SMT and fast and slow tempo limits for each participant was the same as in Experiment 2. After removing outlier ITIs[3], the medians of the ITIs from each block of productions were calculated for each task for each participant. The means of the individual production medians were then calculated and served as the SMT and tempo limits for that participant.

---

criterion). In these cases, the stimulus tempo from the previous iteration was repeated for the next iteration's stimulus.

[3] The criteria for excluding outlier ITIs in the self-paced tapping productions were the same as in Experiment 2. ITIs < 100 ms or ITIs > 1.75*block median were excluded prior to calculating the tempo for the production. There were 15698 ITIs total in the self-paced tapping data. Twenty-three ITIs (< 1%) were removed for being < 100 ms. An additional 490 ITIs (~3%) were removed for being > 1.75*block median. As in Experiment 2, an individual trial was excluded if it included fewer than 10 (out of the target 30) tap intervals. Two blocks (6 additional tap intervals) were excluded based on this criterion.

As in Experiment 2, the self-paced tapping measures were used to calculate normalized produced range (NPR, McAuley et al., 2006), by finding the difference between the slow and fast tempo limits, divided by the participant's SMT (i.e., NPR = (slow limit – fast limit) / SMT).

Serial reproduction (SR) chains. The tempo of each participant's reproduction on each trial was calculated by finding the median of the thirty ITIs of their response. No additional exclusion or preprocessing was performed (other than three exceptions described in Footnote 2).

Drift of serial reproductions. Drift was calculated in the same manner as Experiments 1 and 2 by finding the difference between the tempo of the final reproduction in a chain and the seed tempo. As in the preceding experiments, the slope and x-intercept of a linear regression line fit to drift as a function of seed tempo provided estimates of chain convergence rate and tempo, respectively. Slope values were compared to zero to assess whether or not chains were estimated to converge (i.e., as predicted by the preferred tempo hypothesis).

As in Experiment 2, 99% confidence intervals (CIs) were constructed for regression slopes, and an independent-samples t-test was conducted to test whether slopes differed significantly as a function of whether participants passively observed or synchronized with the auditory stimulus rhythm (i.e., to test the input noise hypothesis).

In order to compare slopes across Experiments 2 and 3, additional t-tests were conducted. Specifically, the slope in the auditory-observe condition of Experiment 2 was compared to the observe condition of Experiment 3, and the slope in the auditory-synchronize condition of Experiment 2 was compared to the synchronize condition of Experiment 3.

Relationship between SR and SMT. As in Experiment 2, a measure of SR tempo relative to preferred tempo was also calculated, by finding the difference between the tempo at each generation and the SMT of the participant making the reproduction and then dividing by the

SMT (i.e., relative tempo = (Stimulus Tempo – SMT) / SMT). The iterative series of relative tempo by chain test the specific prediction that chains would converge to preferred tempo of a participant (which is zero in terms of relative tempo), as measured by their SMT.

In order to assess whether single-participant SR chains converged to the tested participant's preferred tempo, a linear equation was fit to drift as a function of *relative* seed tempo by *participant*. This paralleled the analysis by cohort in Experiment 2. A negative slope indicates that across chains, a participant was biased towards the x-intercept of the function (i.e., the convergence tempo of their chains). The x-intercept tests whether a participant was biased (specifically) towards preferred tempo – an x-intercept of zero (for relative tempo), was expected for participants with a negative slope. The magnitude of the slope indicates how biased the participant was.

Slopes and intercepts for each participant were examined to see whether individuals differed from the on-average slopes and intercepts by condition, and whether there appeared to be a systematic relationship with the width of an individual participant's entrainment region (as measured by NPR).

Results

Spontaneous motor tempo (SMT) and tempo limits. Table 4 summarizes the fast tempo limit, SMT, slow tempo limit, and normalized produced range (NPR) across participants for both experimental conditions. On average, the SMT across participants was 594.3 ms (*SD* = 318.2). The average fast MT was 189.0 ms (SD = 32.9) and the average slow MT was 1530.3 ms (*SD* = 1114.2). Overall, the tempi of the self-paced tapping tasks were in line with previous results (Drake et al., 2000; McAuley et al., 2006), including the results of Experiment 2.

Table 4. Experiment 3: Descriptive summary of SMT and tempo limits by condition

		Auditory	
		Observe	*Synchronize*
Fast	*M*	190.7	187.9
	SD	39.8	27.6
SMT	*M*	575.2	607.6
	SD	346.2	298.3
Slow	*M*	1528.3	1531.7
	SD	1467.8	793.9
NPR	*M*	2.0	2.4
	SD	1.0	1.5

Note. Fast = fast tempo limit. SMT = spontaneous motor tempo. Slow = slow tempo limit. NPR = (Slow – Fast) / SMT. *M* = mean. *SD* = standard deviation. For Fast, SMT, and Slow, units are milliseconds.

Serial reproduction. Analysis of SR tempo data paralleled Experiment 2 in order to evaluate whether the same patterns would hold for single-participant SR chains as were found for multiple-participant SR chains. In Figure 8A, the grey lines show individual chains of tempo reproductions. Seed tempo corresponds to iteration zero, and iteration (serial position in the chain) is shown on the x-axis. Colored lines show the mean tempo averaged across individual chains initiated at the same seed tempo. Consistent with the preferred tempo hypothesis, chains in the observe condition appeared to drift towards a common tempo, with chains initiated by seed tempi faster than average preferred tempo seeming to slow down and chains initiated by seed tempi slower than average preferred tempo seeming to speed up. In contrast, as predicted by the input noise hypothesis, drift towards preferred tempo was reduced in the synchronize condition.
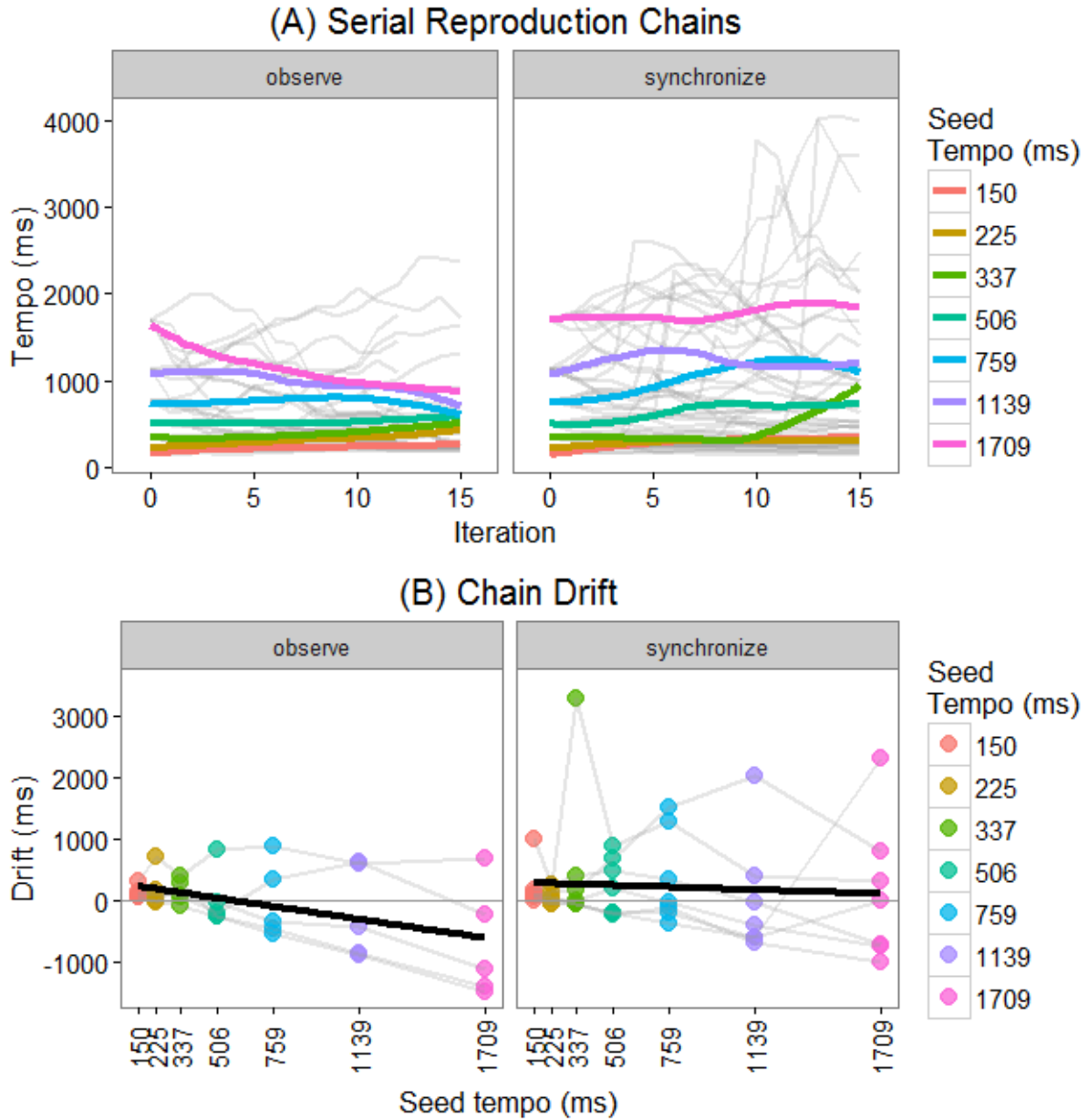
Figure 8. Serial reproduction by condition in Experiment 3
Panel A shows tempo reproductions for each exposure condition (observe, synchronize). Light grey lines plot data from individual chains, and colored lines summarize the mean tempo of reproductions by different participants at each iteration for chains initiated at the same seed tempo. In Panel B, chain drift (i.e., final – seed tempo) is plotted as a function of seed tempo. Each point summarizes drift for an individual chain, with chains belonging to the same participant connected by grey lines. The solid black lines represent the best-fit line to drift as a function of seed tempo across participants in a condition. For lines with negative slope, the intersection of the best-fit line with the x-axis corresponds to the predicted convergence tempo of the chains in that condition.

As in the previous experiments, the change in tempo from seed stimulus to final reproduction for each chain was formalized as drift, which is the signed difference between the final and seed tempo. As shown in Figure 8B, drift was calculated for each SR chain and then a regression line was fit by condition. The results of the regression analysis are shown in Table 5 (left columns under the heading "Drift ~ Seed Tempo"). Consistent with the preferred tempo hypothesis, in the observe condition, slope was negative and significantly different from zero. As expected, the estimated convergence tempo of chains in the observe condition was similar to the average SMT of participants in this condition. Consistent with the input noise hypothesis, the slope in the synchronize condition was not significantly different from zero, although it was also slightly negative. However, an independent samples t-test comparing slopes in the two conditions revealed that the slopes were not significantly different from each other, $t(80) = 1.6$, $p = .12$.

The pattern of results did not differ significantly between Experiment 3 and the parallel conditions in Experiment 2. Comparing slopes in the observe condition of Experiment 3 (which was auditory) to the auditory-observe condition of Experiment 2 revealed that chains converged equally quickly across experiments, $t(66) = 0.4$, $p = .7$. In the synchronize condition of both experiments, slopes also did not differ significantly[4], $t(80) = 1.4$, $p = .2$. The general agreement of drift results across experiments suggests that the SR method is viable for assessing both biases both across participants and within the same participant.

---

[4] However, using the method advocated by Zou (2007) for comparing regression coefficients by constructing the confidence interval of the difference in r's, there was evidence for a reliable difference using a 95% confidence level between the correlation in the synchronize condition in Experiment 3 and the correlation in the parallel

Table 5. Experiment 3: Drift equations by condition

| | Auditory | |
	Observe	Synchronize
Drift ~ Seed Tempo		
Slope	-0.54**	-0.11
CI	-1.00, -0.08	-0.69, 0.47
Y	296.9	287.7
X	551.7	2547.2
R^2	0.24	0.01
Drift ~ Relative Seed Tempo		
Slope	-355.1***	-106.4
Y	91.4	231.1
X	0.3	2.2
R^2	0.64	0.02

Note. CI = lower and upper bounds of the 99% confidence interval for the slope. *Y* = Y-intercept from linear fit. *X* = X-intercept calculated from linear fit. The equation for the "Seed Tempo" fit in the left columns was *Drift (ms) = Slope * Seed Tempo (ms) + Y*. The equation for the "Relative Seed Tempo" fit shown in the right columns was *Drift (ms) = Slope * Relative Seed Tempo + Y*. Asterisks indicate significant difference from zero * $p < .05$, **$p < .01$, ***$p < .001$

Relationship between SR and SMT by condition. The next analysis investigated whether chain drift reflected, specifically, the preferred tempo of individual participants. As in Experiment 2, stimulus tempo at each iteration of the SR chain was converted to "SMT units" by calculating relative tempo (i.e., Relative Tempo = (Stimulus Tempo - SMT) / SMT). In Experiment 3, the SMT used in relative tempo calculations was the SMT of the individual participant who contributed all reproductions to an SR chain (rather than an average across participants in the same cohort, as was done in Experiment 2). In relative tempo units, all participants have a preferred tempo of zero, and one unit of relative tempo is equivalent to a

---

auditory-synchronize condition in Experiment 2. No such reliable difference was found between observe conditions across experiments.

participant's SMT. Stimulus tempi faster than SMT have negative relative tempo; stimulus tempi slower than SMT have positive relative tempo.

Overall, analysis by condition of SR chains and drift using the relative tempo measure replicated the results of the parallel analysis in Experiment 2. Figure 9A shows the change in relative tempo across reproductions by condition. In order of increasing value, the dashed horizontal lines indicate the average fast tempo limit, SMT, and slow tempo limit for both conditions. Drift towards SMT on this plot is reflected by convergence of chains towards zero. As predicted by the preferred tempo hypothesis, chains in the observe condition appeared to drift to zero. Consistent with the input noise hypothesis, chains in the synchronize condition did not appear to converge.

Figure 9B shows the relationship between drift and relative seed tempo by condition. Paralleling the analysis with millisecond stimulus tempi (and of Experiment 2), regression lines were fit by condition as a function of relative seed tempo. The results of the regression analysis are summarized in Table 5 (right columns under the heading "Drift ~ Relative Seed Tempo"). As expected, drift magnitude varied significantly as a negative linear function of relative seed tempo in the observe condition. Consistent with the preferred tempo hypothesis, the estimated convergence tempi of chains was close to preferred tempo as measured by SMT, which, in relative tempo units, was zero. Consistent with the input noise hypothesis, drift towards preferred tempo was reduced in the synchronize condition.
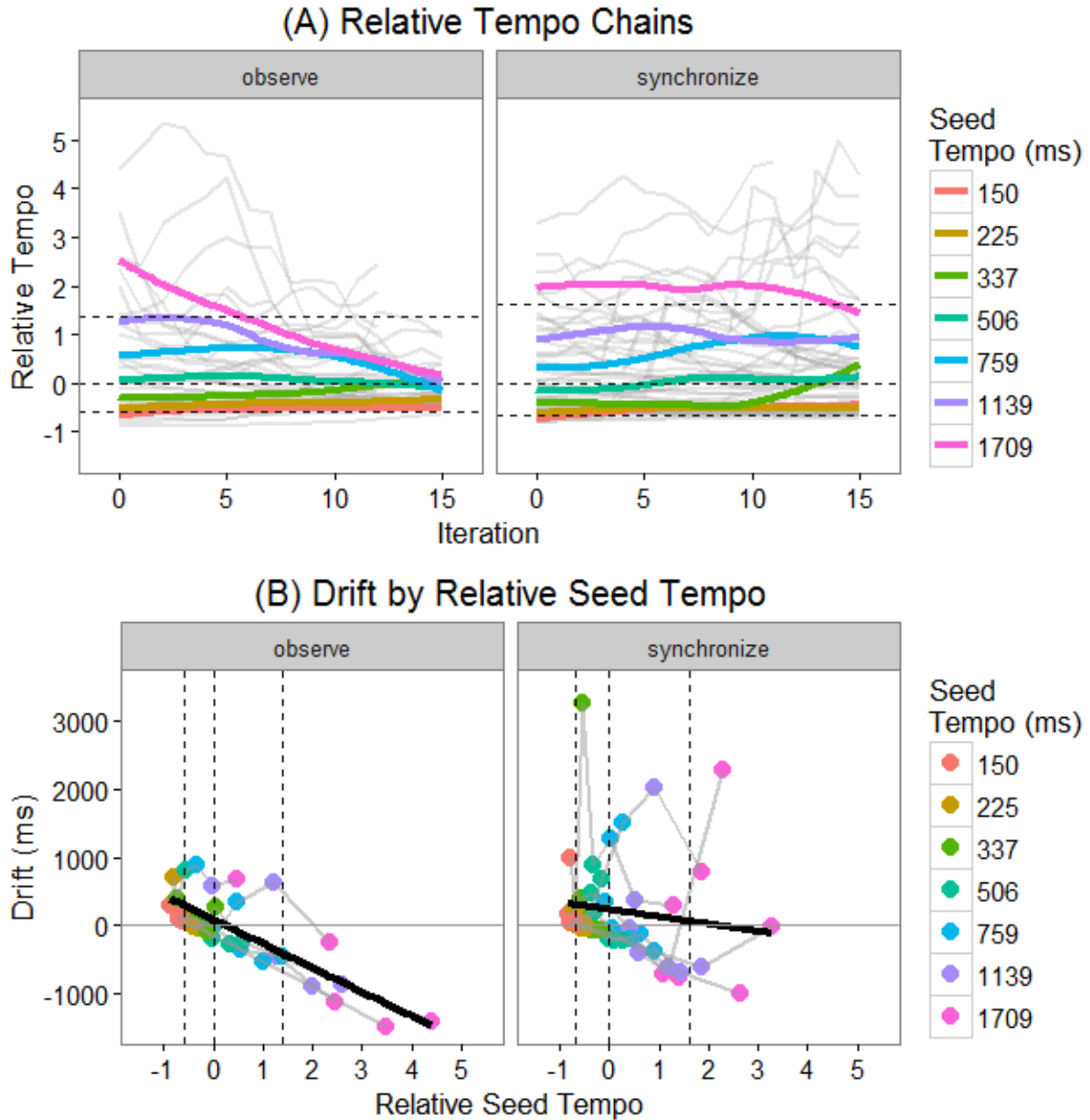
Figure 9. Relative tempo chains and drift in Experiment 3

Panel A shows tempo reproductions (grey lines) in relative tempo units, found by calculating (Stimulus Tempo – SMT) / SMT for each participant, using the participant's SMT. Colored lines summarize over participants for chains initiated by the same seed tempo. Dashed horizontal lines show mean self-paced tapping tempi in SMT units for each condition. The line at y = 0 shows the stimulus tempo equivalent to SMT. Horizontal lines above and below show mean fast and slow (respectively) tempo limits by condition. Panel B shows chain drift by relative seed tempo. From left to right, dashed vertical lines correspond to fast tempo limit, SMT, and slow tempo limit (in SMT units), calculated by condition. Black lines show the best fit lines through the data by condition. For conditions with negative drift slopes, the estimated convergence tempo of chains in each condition is the intersection of this line with the x-axis.

Relationship between SR and SMT by participant. Using single-participant SR chains enabled further decomposition of the results by participant, using differences in SR chain behavior across individuals (relative to the group-level effects) to explore individual differences in tempo bias across participants. Figure 10A shows SR chains in relative tempo units separately for each participant. The left column shows data for participants in the observe condition; the right column shows data for participants in the synchronize condition. Each horizontal panel shows the reproductions of a single participant. On this plot, the colored lines trace the tempo reproductions in a single chain, with color coding for the seed tempo that initiated that chain. The fast and slow tempo limits for the participant (relative to SMT) are shown as the horizontal dashed lines. Figure 10B shows drift as a function of relative seed tempo for each participant. Vertical dashed lines on this plot again show relative tempo limits for the participant whose data are shown in the plot.

Figure 11 plots chain convergence rates (slopes) and convergence tempi (x-intercepts) from the linear fit to drift by relative seed tempo for each participant. Each colored point represents the slope and x-intercept for a single participant. The number label corresponds to the participant ID, and color codes for the participant's NPR. Participants represented by blue points had a wider NPR, whereas participants represented by red points had a narrower NPR. For comparison, the fits by condition are plotted by the empty circles. Shape codes for whether the slope of a participant's drift function was significantly different from zero: triangular points indicate significance at the $p < .01$ level.

For the most part, the convergence rates and tempi obtained for each participant were consistent with the condition fits. Specifically, the relative convergence rate was negative and the convergence tempo was equal or near to preferred tempo (i.e., zero in relative tempo units) for

three out of five of the participants in the observe condition (observe-ID-1, -3, and -5), consistent with the predictions of the preferred tempo hypothesis. From an entrainment perspective, convergence of chains to preferred tempo can be interpreted as the influence of the inherent period of an oscillatory internal timekeeper on the representation of stimulus tempo. This result is less easily explained from an interval perspective.

The relative convergence rate for the other two participants in this condition (observe-ID-2 and -4) were not significantly different from zero, indicating that SR chains of these participants did not demonstrate systematic linear drift. One of these non-conformist participants (observe-ID-2) stands out from the rest of the group as having a slower SMT of about 1161 ms (compared to the group mean SMT of 575 ms) and a wide relative tempo range of 3.17 (compared to the group mean of 2.2). One possibility, from an entrainment perspective, is that the stimulus tempi fell within the entrainment region of this participant's internal oscillatory timekeeper, and thus the timekeeper period could adjust sufficiently to represent the tempo of the stimuli with relatively little bias. The lack of bias by this participant with a relatively wide NPR – despite being in the noisier observe condition – is potentially consistent with this possibility.

The other participant (observe-ID-4) in the observe condition who did not demonstrate systematic drift, however, does not stand out from the group in terms of SMT (510 ms) or relative tempo range (1.71). Although this participant did complete one fewer session than the other participants in this condition (providing only twelve instead of fifteen reproductions per chain), that alone seems unlikely to account for the lack of systematic drift in their chains. The SR chains of observe-ID-4's reproductions shown in Figure 10A provide some insight into how this participant's reproductions may have differed from the others in the condition, leading to the apparent zero-drift result. For this participant, the three chains initiated by the slowest seed

tempi, which were slower than the participant's preferred tempo, appeared to systematically drift towards a slow convergence tempo, whereas the other four chains initiated by seed tempi that were faster than or similar to the participants preferred tempo appeared to systematically drift towards preferred tempo. Thus, one possibility is that this participant had two convergence tempi. This possibility is not accounted for by the current method of fitting a single linear function to drift. At present, it is not entirely clear why this participant's pattern differed from the other participants in the group.

In the synchronize condition, five participants out of seven had slopes that were not significantly different from zero, consistent with the encoding noise hypothesis. Two of these participants (sync-ID-1 and -6) tended towards divergence, as indicated by positive slopes. In line with the result of Experiment 2 (in the auditory-synchronize condition, the two cohorts with non-zero slopes had wider NPR than other cohorts in the same condition), these two participants had the widest NPRs of the group (3.53 and 3.36, compared to the group mean of 2.4), though their SMTs (600 and 522 ms) were similar to the group's (575.2 ms). The SR chains of both of these participants, as shown in Figure 13A, share some similarities to participant observe-ID-4: while chains initiated by seeds similar to or faster than preferred tempo tended to slow down, the slowest seeds were systematically distorted to become even slower. For participant sync-ID-1, this was true for all three slow seeds. For participant sync-ID-6, only the slowest seed was increasingly slowed across generations. The non-zero slopes potentially make sense from an entrainment perspective as overall reduced input noise due to synchronizing with the stimulus rhythm during exposure. The slight divergence of some of the slow chains could result from these tempi approaching the slowest limit of what the participant is capable of considering 'rhythmic'.

71

The other three participants with slopes that did not differ significantly from zero in the synchronize condition (sync-ID-3, -5, and -7) tended towards convergence, as indicated by negative slopes. In contrast to the other participants with non-significant slopes whose chains tended toward divergence, these three participants had relatively narrow NPRs (1.7, 1.8 and 2.1, respectively for sync-ID-3, -5, and -7). Although drift was not significant for these participants, there was a trend towards making biased reproductions, particularly for sync-ID-3 and -5. For these two participants, chains trended towards a convergence tempo slightly slower than (sync-ID-3) or similar to (sync-ID-5) preferred tempo. For sync-ID-7, who had both the widest NPR of this group and the least negative slope, the slowest seed tempo was relatively undistorted across reproductions (consistent with the input noise hypothesis), although chains initiated by the other six seed tempi appeared to converge for this participant. Generally, the tentative relationship between NPR and the sign of drift slopes (i.e., participants with wider NPR had chains that potentially diverged, instead of converging), is consistent with the possibility that input noise is greater for participants with a narrower NPR, leading to tempo bias.

The remaining two participants in the synchronize condition (sync-ID-2 and -4) had negative slopes with convergence tempi equal or near to preferred tempo, suggesting that their reproductions were biased despite synchronizing with the rhythms during exposure. An interesting commonality to note across these two participants was that they both reported using a strategy to keep track of the tempo of the stimuli (either visually tracking the movement of their finger, or imagining a song or beat in their head). These two participants also had more years of musical training (6 and 10 years for sync-ID-2 and -4, respectively) and higher self-ratings of rhythmic ability (both rated themselves at the highest option of 6), compared to the five participants whose chains were not significantly biased towards preferred tempo (only two of

whom had music training and for whom the mean average rhythm rating was 4.4). Although this is a very limited number of data points, it is interesting because it suggests that neither using a strategy nor having musical training appeared to reduce tempo bias in serial reproduction.
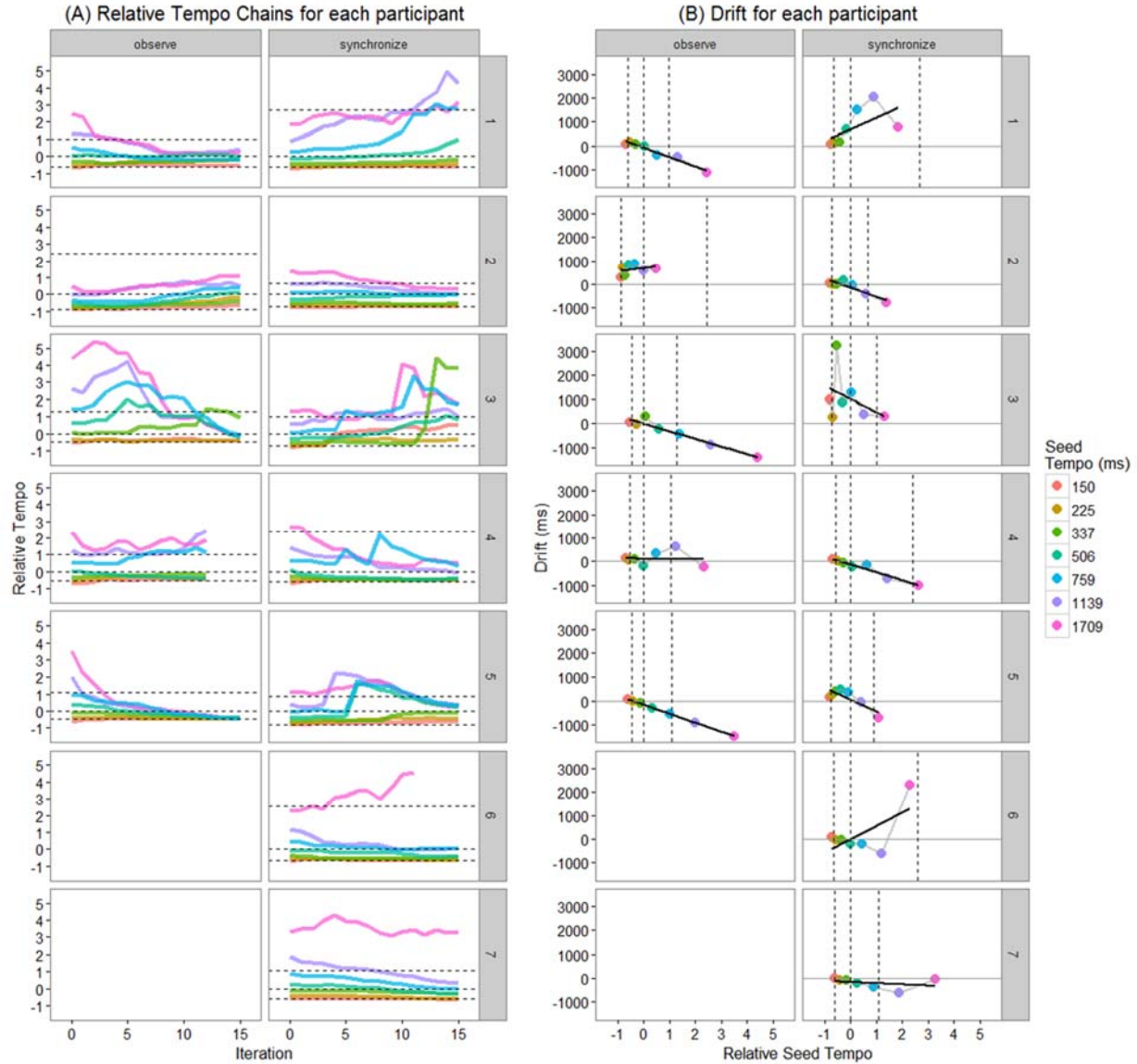
Figure 10. Relative tempo chains and drift by participant in Experiment 3
SR chains in relative tempo units (Panel A) and drift as a function of relative seed tempo (Panel B) for individual participants are shown. Relative tempo = (Stimulus Tempo – SMT) / SMT, using the SMT of the participant. In panel A, each colored line corresponds to one chain of reproductions. Color codes for the seed tempo that initiated the chain. Dashed horizontal lines indicate SMT (at y = 0), and fast and slow tempo limits (relative to SMT) for each participant. In panel B, drift functions are shown for each participant. Seed is coded by color. In this panel, vertical lines indicate the fast and slow tempo limits and SMT of each participant. Black lines indicate linear fit. When chains are biased towards preferred tempo, the black line will have a negative slope and intersect with the origin, indicating that drift is zero when the relative seed tempo is equivalent to SMT.
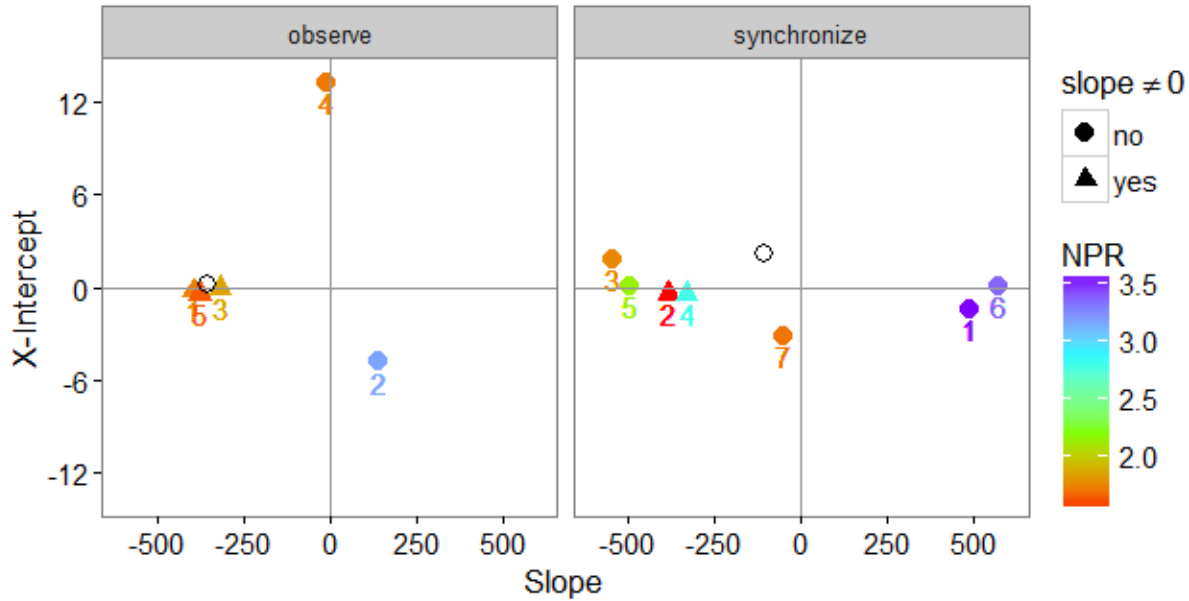
Figure 11. Convergence rates and tempi for relative tempo by participant
Slopes and intercepts from the linear equation of drift as a function of relative seed tempo are plotted in separate panels by condition. Each point is labeled with participant IDs. Color codes for participant NPR (blue for wide range, red for narrow range). Shape codes for whether or not the slope of the equation was significantly different from zero at the $p < .01$ level.

Discussion

Experiment 3 examined the influence of tempo bias on the mental representation of rhythms using single-participant SR chains. As in Experiment 2, preferred tempo and input noise hypotheses were tested by presenting participants with seed sequences that spanned 150 to 1709 ms in tempo, and by manipulating exposure task (*observe* or *synchronize*). Regardless of seed tempo, chains were predicted to converge to preferred tempo, with convergence rates slowed when input noise was low, as expected for chains in the synchronize condition. In contrast to Experiment 2, participants in Experiment 3 were presented with stimuli at the tempi of their own immediately-preceding reproduction. Having only a single participant in the SR chains removed variation across participants in preferred tempo and tempo limits (as well as other potential

factors, such as tempo-tracking strategy use during the experiment) that might have arisen in the previous experiments.

Overall, the analysis of chain drift by condition in Experiment 3 replicated the results of the parallel auditory conditions in Experiment 2.Consistent with the predictions of the preferred tempo hypothesis, the convergence rate (i.e., slope of the drift function) in the observe condition was negative (-0.54), indicating that chains were attracted towards a common tempo. The convergence tempo for the observe condition was 551.7 ms, which was slightly but not much faster than the average preferred tempo for this condition of 575.2 ms (based on SMT). In contrast, SR chains in the synchronize condition did not, on average, demonstrate systematic drift towards a common tempo, as evidenced by an effectively zero convergence rate. Finding systematic drift in the observe but not in the synchronize condition was consistent with the predictions of the input noise hypothesis, and supports the view that moving in synchrony with an auditory rhythm potentially reduces initial encoding noise of the tempo of that rhythm.

Experiment 3 further tested the preferred tempo hypothesis by examining whether SR chains converged specifically to the preferred tempo of the individual participant whose reproductions constituted the chains. The analysis of SR tempo relative to participant SMTs generally supported this prediction. In the observe condition, slopes were negative, indicating that chains converged, and the x-intercept was equivalent to average SMT (zero, in relative tempo units). In the synchronize condition, the average slope was also slightly negative, but not significantly different from zero.

The preferred tempo hypothesis was framed within an entrainment perspective to timing and temporal processing. From this perspective, preferred tempo is the inherent period of an internal oscillatory timekeeper and perceived tempo is represented by the adjusted period of the

oscillator in response to external rhythmic stimulation. Bias towards preferred tempo in a series of reproductions reflects the return of the oscillator's period to its equilibrium state. The finding that participants were biased towards their own preferred tempi in Experiment 3 is generally consistent with this explanation. The reduction in drift in the synchronize condition relative to the observe condition was interpreted from an entrainment perspective, in Experiment 2, as reflecting increased adjustment of the oscillator period in response to the stimulus tempo, based on reduced input noise associated with the stimulus tempo following synchronization. The reduced drift in the synchronize relative to observe condition in Experiment 3 is also consistent with this explanation.

Convergence of chains to the preferred tempi of participants is for the most part inconsistent with an interval perspective on timing. From an interval perspective, tempo is represented as the statistical average of the individual (independently recorded) IOIs of a stimulus rhythm (Drake & Botte, 1993; Wing & Kristofferson, 1973a, 1973b). In general, drift in SR chains is not predicted from this perspective.

However, tempo bias, from an interval perspective, could potentially result from incorporating stimulus statistics across trials in an experiment, akin to adaptation-level effects (Helson, 1964), or central tendency effects (Hollingworth, 1910). Bias towards the mean tempo of stimuli is of particular concern in Experiment 3 since participants had more experience with the stimulus rhythms (fifteen exposures to each chain) than in either Experiment 1 (one exposure) or Experiment 2 (three exposures). If participants were biased by stimulus statistics in Experiment 3, then serial reproductions of tempo should have drifted towards the average of the stimulus tempi, which is 689 ms. Stimulus input noise potentially could influence the relative weighting between memory for the stimulus tempo (on one trial) and memory for the average

tempo across all trials in an experiment. Contrary to an explanation of bias based on stimulus statistics, when chains in the present experiment converged, it was specifically to the preferred tempo of the participant.

A detailed examination of drift slopes and convergence tempi by participant revealed that not all participants demonstrated the identical patterns of drift as others in their same condition. Specifically, three out of five participants in the observe condition had negative slopes and convergence towards preferred tempo (as expected), but the other two participants had slopes near to zero. Likewise, five out of seven participants in the synchronize condition had slopes that were not significantly different from zero (as expected), but the other two participants had negative slopes and chains that converged towards preferred tempo. The variation in input noise effects across individuals can be accounted for within an entrainment perspective by noting that there are individual differences in preferred tempo (oscillator preferred period) and tempo limits (oscillator entrainment region) (Drake et al., 2000; McAuley et al., 2006). In contrast, these kinds of individual differences would not be predicted from an interval or statistical standpoint, since all participants were, at least initially, exposed to stimuli with the same range of tempi.

One potentially intriguing observation was that chains of all six participants with musical training seemed to express bias, and usually towards preferred tempo (however, in two cases, slopes were positive instead of negative, suggesting systematic divergence instead of convergence). In contrast, bias was less consistent across the other six participants who did not have musical training: slopes of these participants were sometimes negative (indicating bias towards a common tempo) and x-intercepts sometimes zero (indicating convergence towards preferred tempo), however this was not always true (for example, two participants in the observe condition had slopes near to zero). However, since participants were not recruited specifically on

the basis of musical training, and most participants with musical training happened to be assigned to the synchronize condition, it is difficult to draw any conclusive meaning from these observations. A future study might recruit musicians and non-musicians in order to better examine the influence of musical training on tempo bias.

One limitation of the present design regards the ability to examine how participants' use of various strategies in order to keep track of stimulus tempi during the experiment might have affected tempo bias. In Experiment 2, several participants reported using a strategy of grouping or subdividing intervals in order to keep track of tempo. Consistent with the possibility that the group length, rather than the ITI, was biased towards preferred tempo, the predicted chain convergence tempo was faster on average than preferred tempo. In contrast, in Experiment 3, no participants reported using this strategy, and, as expected, the predicted chain convergence tempo was closer to the average preferred tempo than in Experiment 2. It is possible that the lack of groupers/sub-dividers underlies the improved agreement between chain convergence tempo and preferred tempo in Experiment 3 relative to Experiment 2. However, it is difficult to draw a conclusion based on the relative paucity of data on this front. A future study might explicitly instruct participants to group or subdivide on some trials, and contrast drift slopes and x-intercepts across trials on which participants receive different instructions or no instructions. This would identify another potential factor that influences bias, namely, grouping strategy.

Conclusion

The results of Experiment 3 using single-participant SR chains generally agreed with the parallel conditions of Experiment 2 using multiple-participant SR chains. In both experiments, chains of auditory rhythms that participants merely observed prior to reproducing converged to preferred tempo. In contrast, chains of auditory rhythms that participants synchronized with

during exposure either did not converge or converged on average more slowly. Together, the results of the two experiments generally support the preferred tempo and input noise hypotheses. Using multiple- vs. single-participant SR chains appeared to highlight different factors that influence tempo bias: removing inter-individual variability by using single-participant chains appeared to improve agreement between estimated chain convergence tempo and SMT in Experiment 3 relative to Experiment 2. However, the effects on tempo bias of input noise, which was manipulated by having participants either passively observe or synchronize with stimulus rhythms during exposure, were less clear-cut with the single-participant chains than with the multiple-participant chains, possibly due to amplification of individual differences (such as differences in preferred tempo or entrainment region width) across serial reproductions.

CHAPTER 5: GENERAL DISCUSSION

This dissertation applied a novel approach to investigating tempo biases in the mental representation of rhythms in the sub-second to second range. The influence of underlying tempo preferences on the representation of rhythms was investigated using a serial reproduction (SR) paradigm. Generally, the serial reproduction paradigm is useful for revealing psychological biases because it magnifies small biases through the process of iteration. Bias in each reproduction leads chains to converge to the underlying expectation that biases reproductions at each iteration. The amount of bias present in each reproduction of an SR chain is proposed to depend on the amount of input noise associated with the stimulus – greater input noise will lead to greater bias, thereby leading SR chains to converge more quickly. Two related hypotheses about tempo bias, namely a preferred tempo hypothesis and an input noise hypothesis were tested in three experiments.

The preferred tempo hypothesis proposed that an inherent tempo preference based on properties of an internal timekeeper with a preferred oscillation period (i.e., tempo) influences the perceptual representation of rhythms, such that the perceived tempo of to-be-learned rhythms are systematically distorted (i.e., biased) to be more similar to the preferred period of the internal oscillator. In the domain of rhythm, a preference for a tempo of about 500 to 600 ms in perceptual preference ratings, spontaneous production, and perceptual discrimination has been observed previously (Baruch et al., 2004; Drake et al., 2000; Drake & Botte, 1993; Fraisse, 1982; McAuley et al., 2006), suggesting that the perception of rhythm is biased towards this tempo. To test the preferred tempo hypothesis, all three experiments in this dissertation varied the seed tempo of isochronous rhythms initiating SR chains. The first participant in a chain was presented with one of these isochronous seed rhythms, and then reproduced it by tapping. Their

reproduction was used to generate the stimulus rhythm for the next participant, and so on, forming a chain of reproductions. Drift across tempo reproductions in SR chains was expected to reveal the influence of inherent tempo biases on the mental representation of rhythm by converging to a common tempo of about 500 to 600 ms.

The input noise hypothesis proposed that increased noise during initial encoding of a stimulus rhythm would lead to greater bias towards preferred tempo during perceptual representation (and consequently, the reproduction) of the stimulus rhythm. In line with past descriptions of serial reproduction for unidimensional stimuli, this generated the prediction that input noise associated with stimulus rhythms would affect the rate at which SR chains converged to preferred tempo. To test the input noise hypothesis, Experiment 2 varied both the modality of stimulus rhythms and the task participants engaged in during exposure to stimulus rhythms. Based on previous work finding general performance advantages for perceiving and reproducing auditory relative to visual rhythms (Collier & Logan, 2000; Grondin & McAuley, 2009; Patel et al., 2005; Repp, 2003; Repp, 2005; Repp & Su, 2013), in the present experiment, presentation of stimulus rhythms in the visual modality was hypothesized to increase input noise relative to the auditory modality, leading chains of visual rhythms to converge more quickly to preferred tempo than chains of auditory rhythms.

In Experiments 2 and 3, during exposure to stimulus rhythms, participants either synchronized finger taps with the rhythms, or passively observed them. Synchronizing with the stimulus rhythms was expected to reduce input noise relative to passively observing them. This prediction was based on several recent proposals related to the role of movement representations in temporal expectation (Arnal, Doelling, Poeppel; Morrillon; Schroeder; Fujioka; Nozaradaran).

82

In the present experiment, the predicted effect of reducing input noise in the synchronize condition was a slowing of chain convergence to preferred tempo.

Preferred tempo hypothesis

Consistent with the preferred tempo hypothesis, across all three experiments, when participants passively observed the stimulus rhythms, chains on average converged to a common tempo, regardless of the seed tempo that initiated the chain. Across experiments, the estimated convergence tempi, which were about 350 to 550 ms, were close to the preferred tempi, as measured by the SMT, of the participants who contributed reproductions, which were about 400 to 600 ms. The preferred tempo hypothesis connects with a broader entrainment-based perspective on timing and temporal processing. From an entrainment perspective, preferred tempo maps onto the inherent period of oscillation of an internal timekeeper. The period of the internal timekeeper transiently adjusts to the period of repetitive external stimulation, but drifts back to its inherent period (in the absence of external input) since this is its equilibrium state (McAuley, 1995; McAuley & Jones, 2003; Large & Jones, 1999; Large & Kolen). Each SR chain is akin to mapping the return of the theoretical internal oscillator period from a perturbed state (the tempo of the seed rhythm) to its equilibrium state (preferred tempo) by taking snapshots of its state (the tempo reproductions) at discrete time steps (each generation). Thus, SR chains of tempo were predicted to drift towards preferred tempo, regardless of seed tempo, because preferred tempo is a stable, equilibrium state of the system.

Specifically, Experiment 1 tested participants in a "portable lab" stationed in a public venue (a restaurant during an art competition). Chains were initiated by isochronous seed rhythms with tempi ranging from 300 to 1200 ms. Each participant listened to the stimulus rhythm, and then reproduced it by tapping, contributing a single reproduction for one chain. The

results of Experiment 1 validated the basic premise of the preferred tempo hypothesis: when participants observed and reproduced rhythms at different tempi, SR chains converged to an estimated tempo of 490 ms, which was similar to though slightly slower than the average SMT across participants, which was 414 ms.

Experiment 2 provided further support for the preferred tempo hypothesis. In this experiment, participants were tested in the laboratory in chains initiated by a wider range of seed tempi ranging from 150 to 1709 ms. Each participant was exposed to stimulus rhythms in seven different chains, providing a single reproduction for seven chains. For conditions paralleling Experiment 1, chains in Experiment 2 also converged. In Experiment 2, the estimated convergence tempo of about 400 ms was faster than the average SMT across participants of about 615 ms. In Experiment 3 – in contrast to Experiments 1 and 2 – each participant provided fifteen reproductions, one at each iteration of seven separate chains. Consistent with the preferred tempo hypothesis, single-participant chains in parallel conditions to Experiments 1 and 2 also converged, specifically to an estimated tempo of about 550 ms. This convergence tempo was similar, but slightly faster than the average SMT across participants, which was 575 ms. Altogether, the convergence of chains across all three experiments to tempi consistent both with a preferred tempo of 500 to 600 ms, as reported in the past literature (Baruch et al., 2004; Drake et al., 2000; Fraisse, 1982; McAuley et al., 2006; Mishima, 1956; Parncutt, 1994; Stern, 1900; van Noorden & Moelants, 1999; Vanneste, 2001; Vierordt, 1868; Wallin, 1911; Woodrow, 1951)), and more specifically to tempi similar to the average preferred tempo of the participants tested (as measured by SMT), supports an entrainment perspective on timing.

In contrast, the convergence of chains to preferred tempo when participants passively observed the stimulus rhythms is generally inconsistent with an interval perspective on timing

(Church, Miller, Meck, & Gibbon, 1991; Gibbon, 1977; Gibbon, Church, & Meck, 1984; Treisman, 1963). From an interval point of view, tempo is represented as the running average of the perceived duration of stimulus IOIs (Drake & Botte, 1993), each of which is centered on the true value of the stimulus, without bias (Wing & Kristofferson, 1973a, 1973b). One widely-cited theory of interval reproduction performance would not predict systematic drift across tempo reproductions (Wing & Kristofferson, 1973a, 1973b). In the Wing & Kristofferson (1973a, b) model, the periodic reproduction of a stimulus duration (e.g., by repeated tapping) is based on the output of a two-component mechanism, which consists of clock and motor processes. The output of the clock specifies the internal representation of the stimulus interval. The mean clock representation is centered on the actual duration of the stimulus, but incorporates noise. The response interval does not correspond precisely to the stimulus or the clock interval both because of noise in the internal clock representation and because of motor delays in producing the response. On average, the motor delay is a constant value, but also includes some random noise. In general, this model predicts that the inter-tap intervals of a reproduction will directly reflect the inter-onset intervals of the stimulus, without bias, because the means of both the stored code for duration and the reproduced interval are centered on the stimulus value. With regards to the present experiment, this model does not predict any systematic error in the produced interval relative to the target stimulus interval. If this were true at the individual level in the present experiment, then there would be no bias to be accumulated across generations into chain-level drift.

One possible way to account for bias from an interval perspective is to consider the influence of stimulus statistics (e.g., the range of stimulus tempi experienced by a participant in an experimental session) on perceived tempo. Past theories related to central tendency effects

(Hollingworth, 1910) or anchoring of psychophysical judgments based on stimulus statistics such as adaptation-level theory (Helson, 1964) or range-frequency effects (Parducci, Perrett, Marsh, 1969) are relevant here. In brief, adaptation-level theory states that participants judge stimulus values relative to "prevailing [internal] norms or adaptation levels" that are based on experimental context (Helson, 1964). Jazayeri & Shadlen (2010) describe a Bayesian model of perceived duration from an internal perspective that incorporates temporal context effects, consistent with adaptation-level theory. In this model, noisy measurements of stimulus duration are combined with prior knowledge based on exposure to stimulus durations throughout the experimental session, during perceptual representation. Consistent with the proposal that stimulus statistics bias perceived duration, participant reproductions of stimulus durations on each trial were systematically biased towards the mean duration of the set of stimulus durations. Although their results are consistent with past theories related to anchoring of psychophysical judgments, Jazayeri & Shadlen (2010) distinguish their model from central-tendency-based explanations in two ways. First, they note that their interpretation offers a mechanistic explanation absent from adaptation-level or range effect explanations. As they say "our work suggests that it is subjects' (implicit) knowledge of their temporal uncertainty that determines the strength of the range effect." They go on to further distinguish their model from past statistical explanations by saying that "the Bayesian account of range effects suggest that production time biases help, rather than harm, subjects' overall performance" (Jazayeri & Shadlen, 2010).

From a bias-by-stimulus-statistics perspective, tempo bias in the present experiments could have resulted from exposure to the range of stimulus tempi in the experiment. The "functional zero" (Helson, 1964) towards which perceptions would be predicted to be biased would be the mean tempo of the stimuli, that is, in accordance with a central tendency effect.

There are, however, several aspects of the results that conflict with such a statistical explanation of tempo bias in the present experiment. First, across all three experiments, the predicted convergence tempi of chains, which ranged from about 350 to 550 ms, were closer to the average of participant SMTs, which were about 450 to 630 ms (as predicted by the preferred tempo hypothesis) than the mean tempo of the stimuli, which was 750 ms in Experiment 1 and was 689 ms in Experiments 2 and 3 (as predicted by central tendency or range explanations). Moreover, in Experiment 1, each participant was exposed to only a single stimulus rhythm, and yet, there was evidence across chain iterations of overall bias towards a tempo of about 500 ms. Finally, in Experiments 2 and 3, drift fits as a function of relative seed tempo (i.e., seed tempo in "SMT units") revealed that estimated chain convergence tempo was, specifically, consistent with participant SMTs, rather than reflecting the mean of stimulus tempi.

More generally, bias in reproduction from memory, and drift towards an intermediate value in SR chains of unidimensional stimuli, may seem like a case of regression to the mean (Galton, 1889; Johnson & Mullally, 1969). The effect of bias is generally that extreme stimulus values are systematically reproduced as less extreme values, consistent with Galton's (1889) early description of the mathematical phenomenon of sons' heights being more moderate than the heights of their fathers. Consistent with this view, in the present set of experiments, when participants observed stimulus rhythms during exposure prior to reproduction, extreme tempi (relative to the tested range) were reproduced as less extreme tempi, and across iterations, bias at each iteration led chains to converge to an intermediate tempo. Drift in SR chains, then, could simply be a case of regression to the mean (Galton, 1889) over iterations.

An argument against bias at each iteration of SR chains in the observe condition as regression to the mean parallels the argument against other central tendency explanations:

although each reproduction in chains in the observe condition was generally a 'more moderate' tempo than the reproduction before, in Experiment 1, participants had no basis for judging what was moderate (because they were exposed to only a single stimulus rhythm). Despite the lack of an experimentally-based frame of reference for evaluating tempo extremity, tempo reproductions in Experiment 1 were biased towards a tempo of about 500 ms. Furthermore, in Experiments 2 and 3, when participants might have had a sense of what an extreme tempo was – based on exposure to the range of stimulus tempi over the course of the experiment – reproductions were, on the whole, biased towards a tempo of about 500 ms, consistent with the preferred tempo hypothesis, rather than being biased towards the mean of the stimulus tempi.

The most fundamental argument against an interpretation of drift in the serial reproduction of rhythm tempo as regression to the mean is that regression to the mean describes the relationship between two *measured* variables. Negative slopes in the linear drift function when participants passively observed the stimulus rhythms (i.e., the relationship between drift and seed tempo shown in Figures 3, 5, and 8) were interpreted as evidence of chain convergence, supporting the preferred tempo hypothesis. However, regression to the mean would generally predict a negative, linear relationship for drift, as well. Problematic to this explanation, though, is the fact that that the predictor variable for the drift function was seed tempo, which represented the true stimulus value, without measurement error.

Huttenlocher et al. (2000) provide a lucid argument in making a similar point to distinguish between bias in reconstruction from memory and regression to the mean. They describe Galton's original formulation of regression to the mean as resulting "because of measurement error…because uncertain values for one measure on a scale…are used to predict uncertain values for a second measure on that scale." In the case where stimulus values are being

used to predict reproduction bias, then "…it has no error. Therefore, error in this measure cannot be invoked to explain why values on the second measure are less extreme than those on the first."

However, an additional point here is that the present results are not wholly inconsistent with a description such as regression to the mean. To again quote Huttenlocher and colleagues: "the term *regression to the mean* can *describe* but does not *explain* the central tendency of judgment unless there is measurement error on both measures." Similarly, in the present set of experiments, the general pattern of results in the observe condition (i.e., drift towards an intermediate, preferred tempo) is not entirely inconsistent with a description such as regression to the mean. However, regression to the mean fails to fully explain the set of results – e.g., why were reproductions biased in Experiment 1? Why did the estimated convergence tempi differ from the mean of the stimulus tempi? To fully explain the results, it is necessary to invoke a mechanistic explanation, as the preferred tempo hypothesis does. Bias in Experiment 1 is proposed to reflect the preferred period of an internal oscillatory timekeeper. In Experiments 2 and 3, the estimated convergence tempi reflect primarily this preferred period, rather than stimulus statistics (though it is possible that both preferred period and stimulus statistics influence tempo reproduction bias).

Finally, as will be discussed further in the following section, an appeal to a statistical explanation of bias in serial reproduction (either central tendency and range effects or regression to the mean) does not fully account for the observed effects on tempo reproduction bias of varying stimulus modality or exposure task.

Input noise hypothesis

The input noise hypothesis was tested by varying stimulus modality (in Experiment 2) and exposure task (in Experiments 2 and 3), in addition to the seed tempo manipulation. The input noise hypothesis was based in part on consideration of the property of serial reproduction that chain convergence rate depends on the ratio of noise in the stimulus input to noise in underlying expectations for the domain of the chain. The motivation for manipulating modality and exposure task within the SR paradigm was to use the predictions of the input noise hypothesis to evaluate current theoretical concerns related to encoding differences across modalities in the rhythm literature and to address the contributions of motor activity to rhythm representations across modalities. The input noise hypothesis predicted that SR chains would converge more quickly to preferred tempo when input noise associated with the stimulus rhythm was large compared to when it was small. Input noise was predicted to be greater for visual compared to auditory rhythms, and for passive observation compared to synchronization during exposure to the stimulus. The outcomes of rhythm modality and exposure task manipulations are discussed in turn in the next two sections.

Rhythm modality. Based on the previous findings of general performance advantages for perceiving and reproducing auditory relative to visual rhythms (Collier & Logan, 2000; Grondin & McAuley, 2009; Patel et al., 2005; Repp, 2003, 2005; Repp & Su, 2013), in the present experiment, input noise was hypothesized to be greater for stimulus rhythms in the visual modality relative to the auditory modality. An increase in input noise for visual relative to auditory rhythms was hypothesized to increase SR chain drift in this condition by nudging individual reproductions at each iteration towards preferred tempo to a greater degree than in the

auditory condition. Thus, it was predicted that the convergence rate would be faster for chains in the visual compared to the auditory condition.

Consistent with predictions, in Experiment 2 convergence rates were faster on average for visual (-0.75) compared to auditory (-0.47) SR chains. Finding faster convergence rates for visual relative to auditory chains suggests that initial encoding of visual rhythms was less reliable than it was for auditory rhythms; and consequently the tempi of visual rhythms were more susceptible to bias by inherent tempo preferences during representation and reproduction. In addition to providing further support for the view that input noise associated with visual rhythms is noisier than it is for auditory rhythms, this result demonstrates that there are unique perceptual biases for tempo across modalities.

In general, a difference in convergence tempo primarily reflects a difference in preferred tempo across modalities, and has implications for theories about preferred tempo and representation across modalities. There was some reason to expect preferred tempo could differ between modalities based on previous findings of tempo differences in auditory versus visual rhythm processing (Collier & Logan, 2000; McAuley & Henry, 2010; Patel et al., 2005; Repp, 2003). For example, Collier and Logan (2000) found that the discrepancy in rhythm discrimination of unimodal auditory or visual rhythms was larger at faster beat periods (125 ms IOIs), and minimal at slower beat periods (500 ms IOIs). Consistent with this result, McAuley and Henry (2010) found that prior presentation of auditory rhythms influenced the interpretation of the rhythmic structure of visual rhythms at a faster 600 ms beat period, but not at a slower, 1,000 ms beat period. In the production domain, participants have been shown to successfully synchronize with auditory rhythms at a tempo as fast as 100-120 ms, but require slowing down to at least 400 ms to synchronize with visual rhythms (Repp, 2003). Thus, based on previous

findings, if preferred tempo differed across modalities, it was expected to be slower for visual than auditory rhythms, which would cause chains in the visual condition to converge to a slower tempo than auditory chains. Contrary to this expectation, in Experiment 2, the convergence tempo of visual chains was faster (352 ms) than the convergence tempo of auditory chains (411 ms). Within the Bayesian framework, the difference in convergence tempi across modalities potentially suggests that prior distributions are modality specific. At present, there does not seem to be an obvious explanation for the unexpectedly faster convergence tempo of visual chains.

Synchronize vs. observe exposure task. Exposure task was manipulated in both Experiments 2 and 3 in order to test the input noise hypothesis. Participants either passively observed the stimulus rhythm (as they did in Experiment 1), or actively synchronized with the stimulus rhythm, prior to reproducing it. The exposure task manipulation addressed the potential role of the motor system in the representation of rhythm. Based on the previous literature, in the present experiments it was expected that synchronizing finger taps with the stimulus during exposure would reduce input noise relative to passively observing the stimulus (Manning & Schutz, 2013; Morillon et al., 2014; Nozaradan et al., 2016). The input noise hypothesis predicted that the effect of this reduction in initial encoding noise in the synchronize condition would be to reduce bias and SR chain drift compared to the observe condition. As with the modality manipulation, the effect of exposure task on input noise (and subsequent representation and reproduction) in a chain of reproductions is most evident in a comparison of chain convergence rates (i.e., drift slope) across conditions. Convergence of chains to preferred tempo was predicted to be reduced (or even eliminated) when bias was reduced in the synchronize condition, leading to flatter slopes in this condition.

Consistent with predictions, on average, the convergence rate of chains in the synchronize condition did not differ significantly from zero in both Experiments 2 and 3, in contrast to convergence rates in the observe condition, which were generally negative and non-zero. Finding reduced convergence in the synchronize vs. observe condition adds support to the proposal that movement benefits accurate representation of temporal information by reducing input noise associated with the rhythm. In addition, a reduction in bias in the synchronize condition suggests that movement may not only increase the accuracy of temporal sensory representations (consistent with past proposals), but further, potentially serves to reduce bias in perceptual representations (Morillon et al., 2016; Nozaradan et al., 2016; Schroeder et al., 2010). That is, beyond 'sharpening' sensory representations, movement additionally shields perceptual representations from systematic bias. This result is consistent with the possibility that increasing the reliability of input representations shifts implicit weighting during perceptual representation from preferred tempo to recent stimulus memory, because moving in synchrony with the stimulus rhythm facilitates veridical initial encoding of that rhythm. Thus, synchronizing in the present experiments seemed to reduce input noise of temporal representations relative to passive observation. The difference in convergence rates across synchronize and observe conditions implies that movement reduced bias in the perceptual representation of rhythms.

Finding a reduction in reproduction bias as a result of synchronizing with the stimulus rhythms during the exposure phase relative to passively observing it contributes some support to current theoretical proposals that the motor system is crucially involved in tracking temporally predictable events (Arnal et al., 2014; Fujioka et al., 2012). This result is consistent with 'active sensing' proposals that neural oscillatory activity associated with the motor system modulates sensory inputs by increasing gain for temporally predictable events, thereby reducing early

93

encoding noise associated with those events (Arnal et al., 2014; Arnal & Giraud, 2012; Morillon & Schroeder, 2015; Morillon et al., 2014; Schroeder et al., 2010). The result that, on average, convergence rates were reduced in the synchronize relative to the observe condition potentially suggests that one effect of reducing early encoding noise is a reduction in representation bias (towards prior expectations).

Interaction of rhythm modality and exposure task. The interaction of rhythm modality and exposure task contributes to a better understanding of encoding differences of temporal information across modalities – particularly to improved understanding of the role of movement in rhythm perception in auditory and visual modalities. In Experiment 2, the results showed that the reduction in bias due to synchronizing with the stimulus rhythms during exposure that was found for auditory rhythms, was reduced for visual rhythms. This suggests that the motor system may make distinct contributions to the encoding of auditory and visual rhythms. And yet, relatively little work to date has addressed this possibility.

Several past studies have noted that the motor system appears to be more responsive to auditory than visual stimulation (Fraisse, 1948; McAuley & Henry, 2010). Much recent neural evidence supports the view that there is meaningful neural overlap between motor-processing and auditory-rhythm processing (Cameron & Grahn, 2014; Chen, Penhune, & Zatorre, 2008; Grahn & Brett, 2007; Grahn & McAuley, 2009; Grahn & Rowe, 2009; Iversen, Repp, & Patel, 2009). Although several propositions have been put forth placing neural oscillatory activity in motor areas at the core of (rhythmic) temporal prediction (Arnal et al., 2014; Fujioka et al., 2012; Morillon et al., 2014; Morillon et al., 2016), the majority of these studies seem to use auditory

stimuli. Thus there seems to be a relative lack of evidence pertaining to motor contributions to encoding of visual rhythms[5].

In one early experiment, Dunlap (1910) found that synchronization of finger taps with a visual flash was generally more error-prone than synchronization with an auditory click. Numerous findings have since corroborated this early result (see Repp, 2005 and Repp & Su, 2013 for reviews). The asymmetrical role of the motor system in rhythm tracking across modalities is further evidenced by findings in the perception domain. For example, although Iordanescu et al. (2013) found a benefit to temporal thresholds of initiating judgment trials by movement for trials presenting auditory stimuli, the same benefit did not accrue for trials presenting otherwise equivalent visual stimuli.

The results of Experiment 2 were consistent with the asymmetrical influence of modality on rhythm perception. In particular, the average convergence slope in the visual synchronize condition, though not significantly different from zero, was still negative (indicating that chains in this condition converged). In contrast, the average convergence slope in the auditory synchronize condition, while also not significantly different from zero, was positive (indicating that chains in this condition actually diverged slightly). Thus there was a greater difference between synchronize and observe slopes in the auditory than the visual modality, consistent with previous findings of general superior performance synchronizing with auditory compared to

---

[5] One exception is a study by (Hove et al., 2013), in which finger tapping synchronization performance and corresponding fMRI response to discrete and continuous pacing signals in auditory and visual modalities were compared, with the general finding that synchronization with continuous – spatially extended – visual signals was better than synchronization with discrete visual signals, but still poorer overall compared to synchronization with auditory signals. In addition, visual signals elicited a reduced response in the basal ganglia (specifically, putamen) relative to discrete auditory stimuli.

visual rhythms (Patel et al., 2005; Repp, 2003, 2005; Repp & Su, 2013), and other studies showing auditory-specific benefits to action (Iordanescu et al., 2013).

Generalization of multiple-participant results to single-participant SR chains

Having separate participants contribute reproductions at each iteration in a chain allows for the possibility that perturbations (such as those based on differences in SMT or in tempo-tracking strategy) may have been introduced in Experiments 1 and 2. Experiment 3 replicated the design of the auditory condition of Experiment 2 with single-participant SR chains, in order to control for variation in individual differences in preferred tempo, tempo limits, and strategy use. Thus, Experiment 3 removed inter-individual variability by having a single participant, presumed to have relatively consistent preferred tempo from session to session contribute all reproductions in an SR chain, with separate groups of participants either synchronizing with or passively observing the stimulus during exposure. In addition to verifying that the results of Experiment 2 would generalize to single-participant chains, Experiment 3 allowed for a closer examination of how properties of an individual's entrainment region (as approximated by SMT and fast and slow tempo limits) might affect tempo bias for rhythms.

Consistent with the preferred tempo hypothesis and the results of Experiment 2, when participants passively observed stimulus rhythms in Experiment 3, chains converged to preferred tempo. Estimated convergence tempo (x-intercept of the drift function) was more consistent with preferred tempo in Experiment 3 than Experiment 2. This was consistent with the possibility that individual participants in Experiment 2 may have introduced perturbations that influenced the tempo of reproductions. Although Experiment 3 did not explicitly control for strategy use, the logic was that an individual participant would be more consistent in the use or non-use of a strategy across reproductions in a chain that separate participants. Since very few participants in

Experiment 3 reported using a strategy to keep track of tempo (and in particular, a strategy of grouping events), the improved convergence of chain convergence tempo with SMTs in Experiment 3 is potentially consistent with the possibility that strategy use in Experiment 2 affected estimated chain convergence tempi.

On average, the effect of synchronizing with the stimulus rhythm reduced the rate of convergence to preferred tempo for single-participant chains in Experiment 3, just as it did for multiple-participant chains in Experiment 2. In terms of validating the use of the SR method for investigating individual differences, this result was promising because it showed that the method is effective for revealing biases regardless of whether reproductions come from different or the same individual.

However, in Experiment 3 the effect of input noise variation due to movement was not as clear-cut as it was in Experiment 2. Slopes and intercepts within condition were more consistent in Experiment 2, where reproductions came from different participants at each generation, than in Experiment 3, in which reproductions came from the same participant at each generation. Specifically, one participant in the observe condition had a chain that did not converge, contrary to expectations. Three participants in the synchronize condition had chains that did converge, contrary to expectations. It is not entirely clear why these participants' data were inconsistent with the input noise hypothesis. This suggests that potentially something about the individual participants (magnified by iteration) caused their chain convergence rates and tempi to vary from expectations. Another possibility is that participants became aware that their reproductions in one block were being used to generate stimuli for the next block, and, in an attempt to finish the experiment more quickly, began tapping more quickly. It is difficult to identify what

characteristics cause this, with so little data. Future studies might include more participants, still using single-participant chains.

Future directions

The present set of results suggest several potential avenues of future work, which include developing a formal model of tempo bias in the perception of rhythm and conducting additional experiments that both test this model, and address several questions raised by the present set of results.

The first potential avenue of future work is to develop a formal Bayesian model to describe (1) how different types of bias (i.e., from preferred tempo versus from exposure to session statistics) combine at the individual level during representation, and (2) how priors at the individual and group level combine to inform drift in SR chains. Bayesian models of serial reproduction posit specific mathematical distributions that correspond to stages of representation and reproduction at each step of an SR chain. Broadly, these stages are input, storage, and output (Xu & Griffiths, 2010). This partitioning of the stages of representation shares similarities with several Bayesian models of perception (e.g., (Jazayeri & Shadlen, 2010; Snyder, Schwiedrzik, Vitela, & Melloni, 2015; Wei & Stocker, 2015). At the input stage, the internal representation of a stimulus is a noisy measurement of the true state of the world. In a Bayesian framework, the input representation is a probability function over possible values of the stimulus, presumed to reach a maximum at the actual stimulus value (i.e., the input is a code for the stimulus plus noise). During storage, noisy inputs are systematically adjusted based on prior knowledge of the world, increasing the accuracy of the representation relative to just sampling from the noisy input distribution. Prior knowledge can be formalized as a specific distribution (e.g., the tempo prior is proposed to take a Gaussian form centered on preferred tempo). In Bayesian terminology, the

prior (preferred tempo) is updated according to Bayes' rule by the data (the noisy stimulus input), generating a posterior distribution. The estimate of the stimulus stored in memory or represented perceptual is a sample or read-out from the posterior distribution. Output, which might consist of a reproduction of the stimulus value or perceptual judgment, may be a read-out of the stored value of the stimulus (Huttenlocher et al., 2000).

There are many details of this general Bayesian approach that could be specified in order to tailor a model to describe the mental representation of rhythms. One possible extension is to incorporate hierarchically-structured prior knowledge (Hemmer & Steyvers, 2009). In the case of tempo, a hierarchical prior at the individual level could capture the process whereby biases from long-term priors (e.g., preferred tempo) combine with transient adjustment to session statistics. At the chain level, a hierarchical prior could describe how variations (in tempo priors) across reproductions (particularly by different participants) combine to yield chain-level patterns of drift. Although previous Bayesian models related to temporal processing have been described (Bååth, 2015; Jazayeri & Shadlen, 2010; Sadakata, Desain, & Honing, 2006; Snyder et al., 2015), little focus has been given to using Bayesian principles to understand how inherent biases for tempo influence the mental representation of rhythms, particularly from an entrainment perspective on timing. Instead, several of these past models adopt an interval perspective on timing (e.g., Jazayeri & Shadlen, 2010), and as such, have focused solely on the influence of session statistics on timing. Thus, the application of a Bayesian approach to tempo biases in rhythm representation would be both a novel extension of the methodology as well as a unique approach to investigating tempo biases.

One empirical follow-up would be to use the SR method with experimentally-manipulated priors (Huttenlocher et al., 2000; Xu & Griffiths, 2010), in a design that parallels

perceptual work that investigates the influence of session statistics on tempo bias (Jones & Mcauley, 2005; McAuley & Miller, 2007). The goal would be to investigate factors that influence how different types of knowledge (e.g., stimulus-based local or global context, inherent preferences) combine in the representation of rhythms. For example, Jazayeri & Shadlen (2010) found that reproductions of stimulus duration were systematically biased by temporal context (i.e., global, experiment context). Specifically, the same stimulus duration embedded in "long", "intermediate", or "short" contexts, in which the mean of the stimulus distributions was a longer, intermediate, or shorter duration, respectively, was systematically lengthened or shortened, depending on context. Most likely, both stimulus statistics and inherent preferences bias the perception of tempo. Thus the prediction for the proposed experiment is that SR chains would converge to preferred tempo, with potential modulation by the range of stimulus tempi. Broadly speaking, such a future experiment would address a fundamental debate in the literature on rhythm and timing, namely, whether the brain's internal timekeeper is more similar to a stop watch, as proposed by interval perspectives on timing, or whether it is more similar to a self-sustaining oscillator, as proposed by entrainment perspectives on timing – or, potentially, whether task conditions (such as movement or stimulus modality) influence the extent to which the internal timekeeper expresses behavior consistent with both.

A formal model of tempo bias could also be combined with future experiments to address lingering issues from the present set of experiments. For example, the estimated convergence tempi of auditory (~400 ms) and visual (~350 ms) chains were faster than SMT (~615 ms) in Experiment 2. This could have potentially resulted from some participants in the chains adopting a grouping strategy. The proposition that grouping strategies may have affected chain convergence tempi was based on reviewing a random selection of participant reports from the

questionnaire at the end of the experiment. Several participants reported either imagining the stimulus events as belonging in groups of two's or three's, or belonging to a short word, phrase or song. For these participants, the length of the group may have been biased towards preferred tempo, rather than the ITI. Participants who used such a grouping strategy could have introduced perturbations into the SR chains (speeding up the tempo of the stimulus rhythms), causing the estimated convergence tempo of the chains to be faster, than if participants did not group events within the stimulus rhythm. A future study could explicitly instruct participants to group in a certain way, and, using the same seed tempi as Experiment 2, evaluate whether grouping strategy influences chain convergence tempo. Instructing participants to group, specifically, in pairs or trios, would generate the specific predictions that chain convergence tempi should be one-half or one-third of preferred tempo as measured by SMT.

Other questions raised by the present set of results, which could be pursued in future work, were that SMTs and slow tempo limits appeared to be slower on average in the synchronize condition compared to the observe condition. This was especially noticeable in Experiment 3, when participants received extensive exposure to the stimuli (compared to Experiments 1 or 2). One possible explanation for the slowing of SMTs or slow tempo limits in Experiment 3 could be adaptation to stimulus statistics across experimental sessions. A future study could explicitly manipulate priors (via stimulus statistics) to explore this possibility. One issue to address would be that stimulus statistics appeared to affect SMT and slow tempo limits in the synchronize condition (when encoding noise is low enough to reduce bias) but not in the observe condition (at least not as much; when encoding noise is high and bias is high).

A related somewhat curious result from the present experiments was that some chains in the Experiments 2 and 3 appeared to diverge, particularly in the synchronize condition. Cohorts

(in Experiment 2) and participants (in Experiment 3) for whom slopes from drift functions were positive (indicating divergence) also tended to have wider NPRs. Although it was expected that a narrower NPR (suggesting a narrower range of rates affording stable entrainment) might lead to increased bias and faster chain convergence towards preferred tempo, wider NPR leading to chain divergence was not entirely expected. There is not a clear-cut explanation for chain divergence at this time, although this result, also, merits future consideration.

Conclusions

This dissertation makes a number of contributions to the field. First, the experiments herein developed a novel method for probing biases in rhythm perception. To date, there have been only two recent studies (Jacoby, 2016; Ravignani, Delgado, & Kirby, 2016) that have used this method to investigate rhythm perception and neither of these have examined tempo biases. Combining the serial reproduction paradigm with previous theoretical approaches to preferred tempo motivated the proposition of two hypotheses, namely the preferred tempo hypothesis, which states that tempo bias in the form of an entrainment region centered on preferred tempo biases the perceived tempo of rhythms, and an input noise hypothesis, which states that the amount of tempo bias in perceptual representations will depend on levels of input noise associated with the stimulus rhythm. These hypotheses provided a novel and useful tool for evaluating the influences of tempo bias on the mental representation of rhythm.

Specifically, the experiments herein showed that serial reproductions of rhythms were biased towards preferred tempo, consistent with an entrainment perspective on timing. Bias towards preferred tempo was greater for serial reproductions of visual than auditory rhythms, consistent with a growing body of evidence that early encoding of tempo is more precise in the auditory compared to visual modality. Moreover, the effect of synchronizing with the stimulus

during exposure appeared to reduce bias for auditory rhythms, but did so to a lesser degree for visual rhythms. Single-participant chains, much like multiple-participant chains, also revealed bias towards preferred tempo, which was reduced when input noise was reduced by synchronizing with the stimulus rhythm.

Overall, by testing the predictions of these two related hypotheses in the domain of rhythm using the method of serial reproduction, the results of this dissertation advance basic research on tempo biases in rhythm perception and production. More broadly, the findings demonstrate the feasibility of using the serial reproduction paradigm in a novel domain and lay the groundwork for the future application of this approach, in both empirical work and computational modelling, to probe underlying biases in the mental representation of rhythm.

REFERENCES

REFERENCES

Arnal, L. H., Doelling, K. B., & Poeppel, D. (2014). Delta-Beta Coupled Oscillations Underlie Temporal Prediction Accuracy. *Cerebral cortex (New York, N.Y. : 1991), 25*(9), 3077-3085. doi: 10.1093/cercor/bhu103

Arnal, L. H., & Giraud, A.-L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences, 16*(7), 390-398.

Bååth, R. (2015). Estimating the distribution of sensorimotor synchronization data: A Bayesian hierarchical modeling approach. *Behavior research methods*, 1-12.

Baese-Berk, M. M., Heffner, C. C., Dilley, L. C., Pitt, M. A., Morrill, T. H., & McAuley, J. D. (2014). Long-term temporal tracking of speech rate affects spoken-word recognition. *Psychological Science, 25*(8), 1546-1553.

Barnes, R., & Jones, M. R. (2000). Expectancy, attention, and time. *Cognitive Psychology, 41*(3), 254-311.

Bartlett, F. C. (1932). Remembering: An experimental and social study. *Cambridge: Cambridge University*.

Baruch, C., Panissal-Vieu, N., & Drake, C. (2004). PREFERRED PERCEPTUAL TEMPO FOR SOUND SEQUENCES: COMPARISON OF ADULTS, CHILDREN, AND INFANTS 1, 2. *Perceptual and motor skills, 98*(1), 325-339.

Becker, M. W., & Rasmussen, I. P. (2007). The rhythm aftereffect: support for time sensitive neurons with broad overlapping tuning curves. *Brain and cognition, 64*(3), 274-281.

Cameron, D. J., & Grahn, J. A. (2014). Neuroscientific investigations of musical rhythm. *Acoustics Australia, 42*(2), 111.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Moving on time: brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *Journal of cognitive neuroscience, 20*(2), 226-239.

Church, R. M., Miller, K. D., Meck, W. H., & Gibbon, J. (1991). Symmetrical and asymmetrical sources of variance in temporal generalization. *Animal Learning & Behavior, 19*(3), 207-214.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*: Routledge.

Collier, G. L., & Logan, G. (2000). Modality differences in short-term memory for rhythms. *Memory & Cognition, 28*(4), 529-538.

Doelling, K. B., Arnal, L. H., Ghitza, O., & Poeppel, D. (2014). Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. *NeuroImage, 85 Pt 2*(2), 761-768. doi: 10.1016/j.neuroimage.2013.06.035

Drake, C., & Botte, M.-C. (1993). Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. *Perception & Psychophysics, 54*(3), 277-286.

Drake, C., Jones, M. R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending. *Cognition, 77*(3), 251-288.

Dunlap, K. (1910). Reaction to rhythmic stimuli with attempt to synchronize. *Psychological Review, 17*(6), 399.

Fraisse, P. (1948). II.-Rythmes auditifs et rythmes visuels. *L'année psychologique, 49*(1), 21-42.

Fraisse, P. (1982). Rhythm and tempo. *The psychology of music, 1*, 149-180.

Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2012). Internalized timing of isochronous sounds is represented in neuromagnetic β oscillations. *The Journal of neuroscience, 32*(5), 1791.

Galton, F. (1889). *Natural inheritance* (Vol. 42): Macmillan and Company.

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review, 84*(3), 279.

Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. *Annals of the New York Academy of sciences, 423*(1), 52-77.

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature neuroscience, 15*(4), 511-517.

Grahn, J. A. (2012). Neural Mechanisms of Rhythm Perception: Current Findings and Future Perspectives. *Topics in cognitive science, 4*(4), 585-606. doi: 10.1111/j.1756-8765.2012.01213.x

Grahn, J. A., & Brett, M. (2007). Rhythm and Beat Perception in Motor Areas of the Brain. *Journal of cognitive neuroscience, 19*(5), 893-906. doi: 10.1162/jocn.2007.19.5.893

Grahn, J. A., Henry, M. J., & McAuley, J. D. (2011). FMRI investigation of cross-modal interactions in beat perception: Audition primes vision, but not vice versa. *NeuroImage, 54*(2), 1231-1243. doi: http://dx.doi.org/10.1016/j.neuroimage.2010.09.033

Grahn, J. A., & McAuley, J. D. (2009). Neural bases of individual differences in beat perception. *NeuroImage, 47*(4), 1894-1903. doi: 10.1016/j.neuroimage.2009.04.039

Grahn, J. A., & Rowe, J. B. (2009). Feeling the Beat: Premotor and Striatal Interactions in Musicians and Nonmusicians during Beat Perception. *The Journal of Neuroscience, 29*(23), 7540-7548. doi: 10.1523/jneurosci.2018-08.2009

Greenberg, S., Carvey, H., Hitchcock, L., & Chang, S. (2003). Temporal properties of spontaneous speech—a syllable-centric perspective. *Journal of Phonetics, 31*(3), 465-485.

Griffiths, T. L., & Kalish, M. L. (2005). *A Bayesian view of language evolution by iterated learning.* Paper presented at the Proceedings of the 27th annual conference of the cognitive science society.

Griffiths, T. L., & Kalish, M. L. (2007). Language evolution by iterated learning with Bayesian agents. *Cognitive Science, 31*(3), 441-480.

Grondin, S., & McAuley, J. D. (2009). Duration discrimination in crossmodal sequences. *Perception, 38*(10), 1542-1559.

Guttman, S. E., Gilroy, L. A., & Blake, R. (2005). Hearing what the eyes see auditory encoding of visual temporal sequences. *Psychological Science, 16*(3), 228-235.

Helson, H. (1964). Adaptation-level theory.

Hemmer, P., & Steyvers, M. (2009). A Bayesian account of reconstructive memory. *Topics in cognitive science, 1*(1), 189-202.

Hollingworth, H. L. (1910). The central tendency of judgment. *The Journal of Philosophy, Psychology and Scientific Methods, 7*(17), 461-469.

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.

Huttenlocher, J., Hedges, L. V., & Vevea, J. L. (2000). Why do categories affect stimulus judgment? *Journal of Experimental Psychology: General, 129*(2), 220.

Iordanescu, L., Grabowecky, M., & Suzuki, S. (2013). Action enhances auditory but not visual temporal sensitivity. *Psychonomic Bulletin & Review, 20*(1), 108-114.

Iversen, J. R., Repp, B. H., & Patel, A. D. (2009). Top-down control of rhythm perception modulates early auditory responses. *Annals of the New York Academy of sciences, 1169*(1), 58-73.

Jackendoff, R., & Lerdahl, F. (1981). Generative music theory and its relation to psychology. *Journal of Music Theory*, 45-90.

Jacoby, N. (2016). Paper presented at the International Conference Music Perception & Cognition, San Francisco, CA, USA.

Jazayeri, M., & Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature neuroscience, 13*(8), 1020-1026.

Johnson, D. M., & Mullally, C. R. (1969). Correlation-and-regression model for category judgments. *Psychological Review, 76*(2), 205.

Jones, M. R. (1976). Time, our lost dimension: toward a new theory of perception, attention, and memory. *Psychological Review, 83*(5), 323.

Jones, M. R., & Boltz, M. G. (1989). Dynamic attending and responses to time. *Psychological Review, 96*(3), 459.

Jones, M. R., Johnston, H. M., & Puente, J. (2006). Effects of auditory pattern structure on anticipatory and reactive attending. *Cognitive Psychology, 53*(1), 59-96.

Jones, M. R., Kidd, G., & Wetzel, R. (1981). Evidence for rhythmic attention. *Journal of Experimental Psychology: Human Perception and Performance, 7*(5), 1059.

Jones, M. R., & Mcauley, J. D. (2005). Time judgments in global temporal contexts. *Perception & Psychophysics, 67*(3), 398-417.

Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science, 13*(4), 313-319.

Kalish, M. L., Griffiths, T. L., & Lewandowsky, S. (2007). Iterated learning: Intergenerational knowledge transmission reveals inductive biases. *Psychonomic Bulletin & Review, 14*(2), 288-294.

Large, E. W., & Jones, M. R. (1999). The Dynamics of Attending: How People Track Time-Varying Events. *Psychological Review, 106*(1), 119-159. doi: 10.1037/0033-295x.106.1.119

Large, E. W., & Kolen, J. F. (1994). Resonance and the perception of musical meter. *Connection science, 6*(2-3), 177-208.

London, J. (2004). *Hearing in time: psychological aspects of musical meter*. Oxford;New York;: Oxford University Press.

London, J., Burger, B., Thompson, M., & Toiviainen, P. (2016). Speed on the dance floor: Auditory and visual cues for musical tempo. *Acta psychologica, 164*, 70-80.

Manning, F., & Schutz, M. (2013). "Moving to the beat" improves timing perception. *Psychonomic Bulletin & Review, 20*(6), 1133-1139.

McAuley, J. D. (1995). Perception of time as phase: Toward an adaptive-oscillator model of rhythmic pattern processing.

McAuley, J. D. (2010). Tempo and rhythm *Music Perception* (pp. 165-199): Springer.

McAuley, J. D., & Henry, M. J. (2010). Modality effects in rhythm processing: Auditory encoding of visual rhythms is neither obligatory nor automatic. *Attention, Perception, & Psychophysics, 72*(5), 1377-1389. doi: 10.3758/app.72.5.1377

McAuley, J. D., & Jones, M. R. (2003). Modeling effects of rhythmic context on perceived duration: a comparison of interval and entrainment approaches to short-interval timing. *Journal of Experimental Psychology: Human Perception and Performance, 29*(6), 1102.

McAuley, J. D., Jones, M. R., Holub, S., Johnston, H. M., & Miller, N. S. (2006). The time of our lives: life span development of timing and event tracking. *Journal of Experimental Psychology: General, 135*(3), 348.

McAuley, J. D., & Miller, N. S. (2007). Picking up the pace: Effects of global temporal context on sensitivity to the tempo of auditory sequences. *Perception & Psychophysics, 69*(5), 709-718.

McAuley, J. D., & Semple, P. (1999). The Effect of Tempo and Musical Experience on Perceived Beat. *Australian Journal of Psychology, 51*(3), 176.

Michaelis, K., Wiener, M., & Thompson, J. C. (2014). Passive listening to preferred motor tempo modulates corticospinal excitability. *Frontiers in human neuroscience, 8*, 252.

Miller, J. E., Carlson, L. A., & McAuley, J. D. (2012). When what you hear influences when you see listening to an auditory rhythm influences the temporal allocation of visual attention. *Psychological Science*, 0956797612446707.

Miller, J. L., Green, K., & Schermer, T. M. (1984). A distinction between the effects of sentential speaking rate and semantic congruity on word identification. *Perception & Psychophysics, 36*(4), 329-337.

Mishima, J. (1956). On the factors of the mental tempo. *Japanese Psychological Research, 1956*(4), 27-38.

Morillon, B., & Schroeder, C. E. (2015). Neuronal oscillations as a mechanistic substrate of auditory temporal prediction. *Ann N Y Acad Sci, 1337*, 26-31. doi: 10.1111/nyas.12629

Morillon, B., Schroeder, C. E., & Wyart, V. (2014). Motor contributions to the temporal precision of auditory attention. *Nature communications, 5*.

Morillon, B., Schroeder, C. E., Wyart, V., & Arnal, L. H. (2016). Temporal Prediction in lieu of Periodic Stimulation. *The Journal of Neuroscience, 36*(8), 2342-2347.

Morrill, T., Baese-Berk, M., Heffner, C., & Dilley, L. (2015). Interactions between distal speech rate, linguistic knowledge, and speech environment. *Psychonomic Bulletin & Review, 22*(5), 1451-1457.

Nozaradan, S., Schönwiesner, M., Caron-Desrochers, L., & Lehmann, A. (2016). Enhanced brainstem and cortical encoding of sound during synchronized movement. *NeuroImage*.

Ollen, J. E. (2006). *A criterion-related validity test of selected indicators of musical sophistication using expert ratings.* The Ohio State University.

Parncutt, R. (1994). A Perceptual Model of Pulse Salience and Metrical Accent in Musical Rhythms. *Music Perception: An Interdisciplinary Journal, 11*(4), 409-464. doi: 10.2307/40285633

Patel, A. D. (2003). Rhythm in language and music. *Annals of the New York Academy of sciences, 999*(1), 140-143.

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

Paternoster, R., Brame, R., Mazerolle, P., & Piquero, A. (1998). Using the correct statistical test for the equality of regression coefficients. *Criminology, 36*(4), 859-866.

Peelle, J. E., & Davis, M. H. (2012). Neural oscillations carry speech rhythm through to comprehension. *Frontiers in Psychology, 3*, 320.

Peterson, R. A. (1984). Asking the age question: A research note. *Public Opinion Quarterly, 48*(1B), 379-383.

Ravignani, A., Delgado, T., & Kirby, S. (2016). *The Cultural Evolution of Structure in Music and Language.* Paper presented at the The Evolution of Language: Proceedings of the 11th International Conference (EVOLANG11).

Recanzone, G. H. (2003). Auditory influences on visual temporal rate perception. *Journal of neurophysiology, 89*(2), 1078-1093.

Repp, B. H. (2003). 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. doi: 10.1080/00222890309603156

Repp, B. H. (2005). Sensorimotor synchronization: a review of the tapping literature. *Psychonomic Bulletin & Review, 12*(6), 969-992.

Repp, B. H., & Penel, A. (2002). Auditory dominance in temporal processing: new evidence from synchronization with simultaneous visual and auditory sequences. *Journal of Experimental Psychology: Human Perception and Performance, 28*(5), 1085.

Repp, B. H., & Penel, A. (2004). Rhythmic movement is attracted more strongly to auditory than to visual rhythms. *Psychological Research, 68*(4), 252-270.

Repp, B. H., & Su, Y.-H. (2013). Sensorimotor synchronization: a review of recent research (2006–2012). *Psychonomic Bulletin & Review, 20*(3), 403-452.

Rimoldi, H. J. (1951). Personal tempo. *The Journal of Abnormal and Social Psychology, 46*(3), 283.

Sadakata, M., Desain, P., & Honing, H. (2006). The Bayesian way to relate rhythm perception and production. *Music Perception: An Interdisciplinary Journal, 23*(3), 269-288.

Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H., & Lakatos, P. (2010). Dynamics of Active Sensing and perceptual selection. *Current opinion in neurobiology, 20*(2), 172-176. doi: 10.1016/j.conb.2010.02.010

Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science, 270*(5234), 303.

Snyder, J. S., Schwiedrzik, C. M., Vitela, A. D., & Melloni, L. (2015). How previous experience shapes perception in different sensory modalities. *Frontiers in human neuroscience, 9*.

Stern, W. (1900). *Über Psychologie der individuellen Differenzen: Ideen zu einer" Differentiellen Psychologie"*: JA Barth.

Su, Y.-H., & Pöppel, E. (2012). Body movement enhances the extraction of temporal structures in auditory sequences. *Psychological Research, 76*(3), 373-382.

Swindle, P. F. (1913). On the inheritance of rhythm. *The American journal of psychology*, 180-203.

Tierney, A., & Kraus, N. (2013). The ability to move to a beat is linked to the consistency of neural responses to sound. *The Journal of Neuroscience, 33*(38), 14981-14988.

Treisman, M. (1963). Temporal discrimination and the indifference interval: Implications for a model of the" internal clock". *Psychological Monographs: General and Applied, 77*(13), 1.

van Noorden, L., & Moelants, D. (1999). Resonance in the Perception of Musical Pulse. *Journal of New Music Research, 28*(1), 43-66. doi: 10.1076/jnmr.28.1.43.3122

Vanneste, V. P., JH Wearden, S. (2001). Temporal control of rhythmic performance: a comparison between young and old adults. *Experimental aging research, 27*(1), 83-102.

Vierordt, K. (1868). *Der zeitsinn nach versuchen*: H. Laupp.

Wallin, J. E. W. (1911). Experimental studies of rhythm and time: II. The preferred length of interval (tempo). *Psychological Review, 18*(3), 202-222. doi: 10.1037/h0071786

Wei, X.-X., & Stocker, A. A. (2015). A Bayesian observer model constrained by efficient coding can explain'anti-Bayesian'percepts. *Nature neuroscience, 18*(10), 1509-1517.

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin, 88*(3), 638.

Wing, A. M., & Kristofferson, A. B. (1973a). Response delays and the timing of discrete motor responses. *Perception & Psychophysics, 14*(1), 5-12.

Wing, A. M., & Kristofferson, A. B. (1973b). The timing of interresponse intervals. *Perception & Psychophysics, 13*(3), 455-460.

Woodrow, H. (1951). Time perception.

Xu, J., & Griffiths, T. L. (2010). A rational analysis of the effects of memory biases on serial reproduction. *Cognitive Psychology, 60*(2), 107-126.

Zou, G. Y. (2007). Toward using confidence intervals to compare correlations. *Psychological methods, 12*(4), 399.