



Published in final edited form as:

*Cognition*. 2015 January ; 134: 232–244. doi:10.1016/j.cognition.2014.10.018.

## Synchronization to auditory and visual rhythms in hearing and deaf individuals

**John R. Iversen,**

Institute for Neural Computation, University of California, San Diego, 9500 Gilman Drive #0559, La Jolla CA 92093

**Aniruddh D. Patel,**

Department of Psychology, Tufts University, 490 Boston Ave., Medford, MA 02155

**Brenda Nicodemus,** and

Department of Interpretation, Gallaudet University, 800 Florida Avenue, NE, Washington, DC 20002, USA

**Karen Emmorey**

Laboratory for Language and Cognitive Neuroscience, San Diego State University, 6495 Alvarado Road, Suite 200, San Diego, California 92120, USA

### Abstract

A striking asymmetry in human sensorimotor processing is that humans synchronize movements to rhythmic sound with far greater precision than to temporally equivalent visual stimuli (e.g., to an auditory vs. a flashing visual metronome). Traditionally, this finding is thought to reflect a fundamental difference in auditory vs. visual processing, i.e., superior temporal processing by the auditory system and/or privileged coupling between the auditory and motor systems. It is unclear whether this asymmetry is an inevitable consequence of brain organization or whether it can be modified (or even eliminated) by stimulus characteristics or by experience. With respect to stimulus characteristics, we found that a moving, colliding visual stimulus (a silent image of a bouncing ball with a distinct collision point on the floor) was able to drive synchronization nearly as accurately as sound in hearing participants. To study the role of experience, we compared synchronization to flashing metronomes in hearing and profoundly deaf individuals. Deaf individuals performed better than hearing individuals when synchronizing with visual flashes, suggesting that cross-modal plasticity enhances the ability to synchronize with temporally discrete visual stimuli. Furthermore, when deaf (but not hearing) individuals synchronized with the bouncing ball, their tapping patterns suggest that visual timing may access higher-order beat perception mechanisms for deaf individuals. These results indicate that the auditory advantage in

---

© 2014 Elsevier B.V. All rights reserved.

Corresponding Author: John R. Iversen, Institute for Neural Computation, University of California, San Diego, 9500 Gilman Drive # 0559, La Jolla CA 92093-0559, tel: + 1 (619) 757-6210, jiversen@ucsd.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

rhythmic synchronization is more experience- and stimulus-dependent than has been previously reported.

## Keywords

timing; synchronization; audition; vision; deafness

---

## 1. Introduction

Rhythmic synchronization (the ability to entrain one's movements to a perceived periodic stimulus, such as a metronome) is a widespread human ability that has been studied for over a century in the cognitive sciences (Repp & Su, 2012). Across many studies, a basic finding which has been extensively replicated is that entrainment is more accurate to auditory than to visual rhythmic stimuli with identical timing characteristics (e.g., to a metronomic series of tones vs. flashes). Interestingly, when nonhuman primates (Rhesus monkeys) are trained to tap to a metronome, they do not show this modality asymmetry, and furthermore they tap a few hundred ms after each metronome event, unlike humans who anticipate the events and tap in coincidence with them (Zarco, Merchant, Prado, & Mendez, 2009). Thus, the modality asymmetry (and anticipatory behavior) seen in human synchronization studies is not a generic consequence of having a primate brain, and instead reveals something specific to human cognition.

In cognitive research, the auditory-visual asymmetry in human rhythmic synchronization accuracy is often taken as reflecting superior temporal processing within the auditory system. In the current study we asked if this asymmetry is indeed an inevitable consequence of human brain function, or if the asymmetry can be tempered (or even eliminated) by altering the characteristics of visual stimuli. In addition, we investigated if developmental experience plays a role in the accuracy of visual timing abilities and rhythmic perception by studying profoundly and congenitally deaf individuals.

### 1.2 Modality differences in temporal processing

The question of modality differences in the perception of time and rhythm has been a matter of debate for at least a century, and the question is of interest because it constrains theories of the cognitive architecture of timing. Is there an amodal timing center that can be accessed through the various senses, or is timing inherently tied to audition? Explanations have ranged from early claims that purely visual rhythm exists on an equal footing with auditory rhythm (e.g., Miner, 1903), to claims that any rhythmic sense one gets from a visual stimulus is due to an internal recoding into auditory imagery ('auralization', e.g., Guttman, Gilroy, & Blake, 2005). The former view is consistent with a more general timing facility, while the latter obviously ties timing uniquely to the auditory system.

Most existing research, and common experience, appear to support a significant deficit in rhythmic timing behavior driven by visual compared to auditory input. Evidence to date indicates that no discrete rhythmic visual stimulus can elicit the precise motor synchronization that humans exhibit to auditory rhythms (Repp & Penel, 2004). For the

most part this view has been based on comparing synchronization with tones to synchronization with discretely timed visual stimuli (typically flashes), which are similar to tones in having abrupt onsets and a brief duration. When comparing tones and flashes, one observes significantly poorer synchronization with visual flashes (Repp, 2003), a dominance for sound in audio-visual synchronization paradigms (Repp & Penel, 2004), poorer recall of visual rhythmic patterns, despite training (Collier & Logan, 2000; Gault & Goodfellow, 1938), and temporal illusions in which the timing of an auditory stimulus influences perception of the timing of a visual stimulus (reviewed in Recanzone, 2003). The consistent superiority of audition in these studies has supported the idea that the auditory system is specialized for the processing of time, as expressed in the dichotomy “audition is for time; vision for space”, a generally-accepted tenet that has guided much thought and experiment about the relationships between the modalities (Handel, 1988; Kubovy, 1988; Miner, 1903).

Poorer synchronization to flashing vs. auditory metronomes is surprising because the individual flashes generally occur at rates well within the temporal precision of the visual system (Holcombe, 2009). Although the flashes within visual rhythms are individually perceived, they do not give rise to the same sense of timing as do temporally equivalent auditory stimuli and thus cannot as effectively drive synchronization (Repp, 2003). This result has led to the suggestion that there is a fundamentally different mode of coupling between rhythmic information and the motor system, depending on whether the information is auditory or visual (Patel, Iversen, Chen, & Repp, 2005). This notion is supported by behavioral studies suggesting that flashes on their own do not give rise to a strong sense of beat (McAuley & Henry, 2010), and findings based on autocorrelation measures of successive inter-tap intervals suggest that the timing of taps to auditory and visual stimuli may be controlled differently (Chen, Repp, & Patel, 2002). Neuroimaging studies have also shown that synchronization with flashing visual stimuli depends on different brain regions than synchronization with temporally equivalent rhythmic auditory stimuli, including much stronger activation (during auditory-motor vs. visual-motor metronome synchronization) in the putamen, a region of the basal ganglia important in timing of discrete intervals (Grahn, Henry, & McAuley, 2011; Hove, Fairhurst, Kotz, & Keller, 2013; Jäncke, Loose, Lutz, Specht, & Shah, 2000; Merchant, Harrington, & Meck, 2013).

### 1.3 Temporal processing of moving visual stimuli

Although tapping to flashes lacks the precision of tapping to tones, there is ample evidence of precise visuomotor timing in other kinds of tasks. For example, the brain accomplishes precise visuomotor timing when moving objects and collisions are involved, e.g., when catching a ball or hitting a ball with a tennis racket (Bootsma & van Wieringen, 1990; Regan, 1992). Such visuomotor acts are typically not periodic, but it is possible that the accuracy demonstrated in such acts could occur when synchronizing with a moving visual stimulus with a periodic “collision point.” A critical question, then, is how well synchronization can be driven by a moving visual stimulus with a periodic collision point. To address this question, we investigated visuomotor synchronization with a simulated bouncing ball that collided periodically (and silently) with the floor.

It is important to note that the synchronization behavior used in this study involves discrete movements (i.e., finger taps) aligned with discrete points in time (collision points), and is thus distinct from the large body of work demonstrating precise visuomotor timing in *continuous* movement synchronization with moving visual stimuli (e.g., Amazeen, Schmidt, & Turvey, 1995). While participants show a high degree of temporal precision in such continuous synchronization tasks, such synchronization is thought to employ a fundamentally different implicit timing mechanism than that involved in driving discrete, explicitly timed, rhythmic responses, which require explicit representation of temporal events (Huys, Studenka, Rheaume, Zelaznik, & Jirsa, 2008; Zelaznik, Spencer, & Ivry, 2002). Thus, we are asking how well discrete, repeated moments in time can be coordinated with moving visual stimuli.

There has been little research examining discrete rhythmic tapping to periodically moving visual stimuli. Recently, Hove et al. (2010) demonstrated that synchronization with moving visual stimuli (images of vertical bars or fingers moving up and down at a constant speed) was considerably better than to flashes, although it was still inferior to synchronization to an auditory metronome. The current work extends this finding in two directions. First, we ask if presenting more physically realistic motion trajectories with a potentially more distinct temporal target could enhance visual synchronization to the same level as auditory synchronization. Such a result would be noteworthy since to date no purely visual stimulus has been shown to drive rhythmic synchronization as well as auditory tones.

Second, the current work takes a different approach to testing synchronization than much of the past work. In contrast to previous studies, which used the failure of synchronization at fast tempi to define synchronization ability (Hove et al., 2010; Repp, 2003), we chose a slower tempo (at which synchronization would be expected to be successful for all stimuli) because we wanted to study the temporal quality of *successful* synchronization in order to gain insight into the mechanisms of synchronization. That is, rather than pushing the system to failure, we instead characterize a range of measures of dynamic quality of synchronization, recognizing that 'successful' synchronization can encompass a wide range of behaviors. We thus chose a tempo of 100 beats per minute (600 ms inter-onset-interval; IOI), which is a tempo where the majority of participants are expected to synchronize successfully with both tones and flashes (Repp, 2003). Ultimately, our aim was to determine if synchronization to a moving visual stimulus could match the precision seen for auditory stimuli, and if it demonstrates other temporal properties in common with tapping to sound (e.g. as measured by autocorrelation of successive taps), which would suggest the possibility of a common mechanism. If so, this would challenge the view that the auditory system is uniformly more effective at driving periodic discrete movement.

#### 1.4 Influence of deafness on visual processing

Can the ability to synchronize with visual stimuli be modified by experience? Attempts to improve visual temporal processing in hearing participants through training have generally not been successful (Collier & Logan, 2000). However, the radically different experience with sound and vision experienced by adults who are profoundly deaf from birth could have a significant impact on the ability to synchronize with visual stimuli. Deafness (often in

conjunction with sign language experience) has been shown to affect a number of visual and spatial abilities (see Bavelier, Dye, & Hauser, 2006; Pavani & Bottari, 2012, for recent reviews); however, the extant literature is unclear with respect to whether deafness might be predicted to improve or impair visual synchronization ability.

On the one hand, auditory deprivation has been shown to enhance aspects of visual processing, e.g., visual motion detection (Bavelier et al., 2001), attention to the visual periphery (Proksch & Bavelier, 2002), and speed of reaction to visual stimuli (Bottari, Caclin, Giard, & Pavani, 2011). In particular, deaf individuals are faster and more accurate at detecting the direction of motion of a stimulus presented in the visual periphery (e.g., Neville & Lawson, 1987), although deaf and hearing groups do not differ in coherent motion discrimination thresholds (e.g., Bosworth & Dobkins, 2002). Deaf individuals are also faster to detect abrupt visual onsets (Bottari et al., 2011; Bottari, Nava, Ley, & Pavani, 2010; Loke & Song, 1991). Several researchers have proposed that enhanced visual abilities in deaf individuals stem from selective changes in the “motion pathway” along dorsal visual stream (e.g., Armstrong, Neville, Hillyard, & Mitchell, 2002; Bavelier et al., 2006; Neville & Bavelier, 2002). Recent evidence has also suggested that behavioral changes in visual processing are associated with altered neural responses in striate cortex (Bottari et al., 2011) as well as in primary auditory cortex (Scott, Karns, Dow, Stevens, & Neville, 2014). It is possible that greater sensitivity to motion direction and to visual onsets could improve visuomotor rhythmic synchronization to a moving and/or a flashing visual stimulus, perhaps to the level of auditory synchronization observed for hearing individuals. Such a result would indicate that a strong asymmetry between visual and auditory synchronization may not be an inevitable consequence of human neural organization.

On the other hand, several studies have also documented a range of temporal perception deficits in deaf individuals. Kowalska and Szlag (2006) found that congenitally deaf adolescents were less able to estimate temporal duration, and Bolognini et al. (2011) found that deaf adults were impaired compared to hearing adults in discriminating the temporal duration of touches. Bolognini et al. (2011) also provided TMS evidence that auditory association cortex was involved in temporal processing at an earlier stage for deaf participants, which correlated with their tactile temporal impairment. Finally, several studies have documented temporal sequencing deficits in deaf children (e.g. Conway et al., 2011; Conway, Pisoni, & Kronenberger, 2009). These findings have led to the suggestion that auditory experience is necessary for the development of timing abilities in other modalities. In particular, Conway et al. (2009) proposed the “auditory scaffolding” hypothesis, which argues that sound provides a supporting framework for time and sequencing behavior. Under conditions of auditory deprivation, this scaffolding is absent, which results in disturbances in non-auditory abilities related to temporal or serial order information. If visual synchronization were similarly bootstrapped from auditory experience, then one would expect poorer rhythmic visual synchronization in deaf individuals than in hearing individuals.

## 1.5 Overview of experiments, dependent measures, and hypotheses

To investigate these questions, we measured sensorimotor synchronization by deaf and hearing individuals using a variety of isochronous stimuli: a flashing light (presented in central and peripheral visual fields), a simulated bouncing ball, and auditory tones (for hearing participants only). As mentioned above, deaf individuals have been shown to allocate more attention to the visual periphery compared to hearing individuals, possibly because they rely on vision to monitor the edges of their environment (Bavelier, Dye, & Hauser, 2006). In addition, auditory cortex exhibits a larger response to peripheral than to perifoveal visual stimuli for deaf individuals (Scott et al. 2014). To examine whether heightened visual attention to the periphery (and the possible greater involvement of auditory cortex) might facilitate visuomotor synchronization, the flash stimuli were presented to the peripheral visual field, as well as centrally (on different trials). We hypothesized that group differences in visuomotor synchronization with flashes might reveal an advantage for deaf participants and that this might be most evident for stimuli presented to the visual periphery,

For all stimulus types, synchronization was measured by asking participants to tap a finger in time with the stimuli. Several measures of performance were quantified: 1) variation in tap timing, which addresses how effectively the timing of the stimulus is represented and is able to drive motor output, 2) mean asynchrony (deviation of tap time from stimulus target time) as a measure of the accuracy of timing, 3) the prevalence of outright failures to synchronize, and 4) autocorrelation structure of inter tap interval (ITI) time series. These measures will provide insight into possible presence and nature of timing mechanisms, such as error correction and metrical time keeping. While measures 1–3 are measures of the quality of synchronization, autocorrelation provides a way to make inferences about underlying timing mechanisms, because it describes how the timing of behavior depends on past history. Similar autocorrelation patterns are consistent with similarities in underlying timing mechanisms (Wing, 2002), and autocorrelation can provide evidence for more complex hierarchical timing processes central to music perception (Vorberg & Hambuch, 1978), as explained in greater detail in section 2.4 below. We hypothesized that measures 1–3 would reveal that tapping in time to a moving visual stimulus is superior to tapping to a flashing visual stimulus and that the level of performance is comparable to tapping to tones. For autocorrelation, we predicted an equivalent degree of error correction when tapping with moving visual stimuli and tones, as revealed by negative lag-1 autocorrelation values. However, higher-order autocorrelations, at lags  $> 1$  (which provide evidence for more complex timing processes) might only be present for synchronization to tones, which would suggest that timing processes involved in the hierarchical perception of time may be limited to the auditory domain, despite equivalent performance across modalities. (Details about specific predictions concerning autocorrelation results at different lags are presented in section 2.4.)

## 2. Material and Methods

### 2.1. Participants

Twenty-three deaf signers (11 women;  $M_{\text{age}} = 28.9$  years) and twenty-two hearing non-signers (15 women;  $M_{\text{age}} = 27.0$  years) participated. The two groups did not differ in age,  $t(43) = .440$ . The deaf and hearing groups also did not differ in their level of education (a reasonable proxy for socioeconomic status). The majority of participants in both groups were college educated, and the mean number of years of education did not differ between the deaf participants (14.9 years;  $SD = 2.0$ ) and the hearing participants (15.5 years;  $SD = 1.7$ ),  $t(43) = .265$ . The deaf participants were born deaf ( $N = 20$ ) or became deaf before age three years ( $N = 3$ ) and used American Sign Language (ASL) as their primary language. All deaf participants were either severely deaf ( $> 70\text{dB}$  loss;  $N = 3$ ) or profoundly deaf ( $> 90\text{ dB}$  loss;  $N = 20$ ). Hearing participants reported normal hearing and had no knowledge of ASL (beyond occasional familiarity with the fingerspelled alphabet). Hearing participants were recruited from the San Diego area and tested at San Diego State University (SDSU) and were not selected on the basis of musical experience. Deaf participants were either tested at SDSU or at a deaf convention in San Diego. All participants were paid for their participation. This study was approved by the Institutional Review Board of San Diego State University.

### 2.2. Stimuli

Three isochronous rhythmic stimuli were used in this study: A flashing visual square, an animated bouncing ball, and auditory tones. The inter-event interval was 600 ms. Stimulus details were as follows. *Flash*: A white square ( $2.7^\circ$  wide) flashed on for 100 ms on a black screen (Dell Latitude e600, LCD, 60Hz refresh rate, response time  $\sim 16$  ms (one frame) for transition from black to white, flash duration was six frames). The flashing square was located at the center (3 trials), or at one of four randomly chosen corners of the screen ( $10.7^\circ$  eccentricity; position constant during each trial; 3 trials). Fixation was maintained on a cross at the center of the screen. *Bouncing Ball*: A silent, realistic image of a basketball ( $2.7^\circ$  diameter) moved vertically with its height following a rectified sinusoidal trajectory resulting in a sharp rebound (vertical distance moved was approximately 2 ball diameters). A stationary white bar at the bounce point yielded the illusion of an elastic collision of the ball with the floor. *Tones*: Sinusoidal 500 Hz tone bursts with 50 ms duration, 10 ms rise/fall presented at a comfortable level over headphones.

### 2.3. Procedure

Participants sat 25" from an LCD display, with a MIDI drum pad in their lap (Roland SPD-6). Instructions and stimuli were presented using Presentation (Neurobehavioral Systems), which also collected responses from the drum pad through a MIDI to serial interface (Midiman PC/S). For hearing participants, all stimuli were used, and instructions were presented as text. For deaf participants, only the visual stimuli were presented, and videotaped instructions were presented in ASL on the computer screen. ASL was the first language for the deaf participants, and therefore both groups received instructions in their native language. The instructions were to tap in synchrony with the stimulus at a steady tempo using the dominant hand. The tap target was the flash or tone onset or the moment of

rebound of the ball. Stimuli were presented blocked in one of four random orders for each participant group, counterbalanced across participants. Each block contained instructions, a demonstration of each stimulus and three identical experimental trials (6 for the Flash condition because there were two flash positions). Data from the trials for a given stimulus were pooled. Stimuli for each trial lasted 30 seconds. The overall experiment time averaged 18 minutes for deaf participants and 20 minutes for hearing participants. While hearing participants had more trials (because of the inclusion of auditory stimuli) the video instructions to the deaf participants took longer, which reduced the difference in total time for completing the experiment.

## 2.4. Analyses

Tapping time series were analyzed using circular statistics (Fisher, 1993). Tap times were converted to relative phase (RP), a normalized measure of the time of the tap relative to the nearest stimulus onset, using the stimulus onset times of the immediately preceding and following stimuli to define a reference interval (see Figure 1E below). RP ranges from  $-0.5$  to  $+0.5$ : a RP of zero indicates perfect temporal alignment between tap and stimulus, a negative RP indicates that the tap preceded the stimulus, and positive RP indicates that the tap followed the stimulus. RP values can be represented as vectors on the unit circle (Figure 1B, F). Synchronization performance was quantified in several ways. A first measure is the circular standard deviation ( $sd_{\text{circ}}$ ) of the tap RPs, defined as  $\sqrt{-2\ln(R)}$ , where  $R$  is the magnitude of the sum of RP vectors (Fisher, 1993). Taps consistently timed relative to the stimulus have relative phases tightly clustered around a mean value, and thus a low  $sd_{\text{circ}}$ . A second measure was the mean value of the tap RPs, which indicates how close taps were on average to the target, and which is the circular analog to the standard measure of tap-to-stimulus asynchrony (Repp, 2005). Raw values are in cycles, but can be converted to ms by multiplying by the stimulus period of 600 ms. Parametric statistical analysis was performed, except in cases when one or more variables violated normality, in which case non-parametric statistical tests were used.

In addition to these measures of accuracy, a third measure characterizing the relation of successive taps was examined by calculating the autocorrelation of the inter-tap-interval time series (ITI; timeseries: Figure 1C, G; autocorrelation: Figure 1D, H). Autocorrelation is often used to make inferences about underlying timing mechanisms because it describes how the timing of behavior depends on past history. Similar autocorrelation patterns are consistent with similarities in underlying timing mechanisms (Wing, 2002). Tapping with auditory metronomes typically leads to negative values of the lag-1 autocorrelation (AC-1), meaning that a longer interval tends to be followed by a shorter one, which is often taken as a sign of online error correction. Conversely, positive AC-1 values are indicative of slower drifts in the ITI (e.g. relatively long ITIs are followed by long ITIs). Most commonly, models of simple synchronization predict that higher-order (lag-2, lag-3 etc.) autocorrelations are zero, meaning that the timing of each tap depends only on the duration of the preceding interval (Repp, 2005; Wing, 2002). However, humans are capable of more complex synchronization at multiple periodic levels, so-called metrical timekeeping, an ability that is central to music perception and production (e.g. Patel, et al., 2005). Non-zero higher-lag autocorrelations have been proposed as signs of this metrical time keeping,



specifically interval timing that operates at multiple nested levels (Keller & Repp, 2004; Vorberg & Hambuch, 1978). Relevant to the current work, existing evidence suggests that *specific* patterns of higher-order autocorrelations characterize metrical timekeeping. In particular, a relatively negative lag-3 autocorrelation (AC-3), with other high-order lags being close to zero, has been found experimentally when participants were instructed to synchronize their response in groups of 2, or binary meter (Vorberg & Hambuch, 1978). Grouping by three or four, in contrast, is characterized by positive lag-3 and lag-4 autocorrelations respectively, with other high-order lags zero. This is an empirical result, and while the mechanistic reason for these patterns is not yet understood, they provide a template for searching the present results for evidence of metrical timekeeping.

Based on our goal of characterizing differences in the quality of *successful* synchronization, we used a strict criterion to include only strongly synchronized trials. Trials without strong synchronization (Rayleigh test,  $P > 0.001$ ) were excluded from analysis of these three measures ( $sd_{\text{circ}}$ , mean RP, and autocorrelation). Using this exclusion criterion only the Flash condition had an appreciable number of excluded trials (average  $\sim 31\%$ ). All other conditions combined had on average 0.8% excluded trials (and these were caused by only one or two individuals). Additionally, for the autocorrelation analysis of ITI time series any trials with missing taps were excluded ( $\sim 5\%$ ). The first four taps of each trial were omitted, to avoid startup transients, leaving on average  $45.3 \pm 3.8$  analyzed taps per trial. Finally, in a separate analysis, the percentage of failed trials was also calculated, by counting the number of trials that failed the Rayleigh test, using a conservative threshold of Rayleigh  $P > 0.1$ .

### 3. Results

#### 3.1. Modality differences in synchronization

To establish a baseline for the comparisons of interest, we replicated previous results concerning synchronization with discrete auditory and visual isochronous stimuli in hearing participants. Figure 1 shows representative examples of tapping to tones and flashes by a hearing individual and illustrates the main quantitative measures used to characterize synchronization in this study. The top row is a typical example of tapping to an auditory metronome, showing stable relative phase (RP) that is close to zero (meaning that taps are close to the sound onsets) and correspondingly small variations in inter-tap interval (ITI; see Figures 1B and 1C, respectively). The sequential patterning of ITI is measured by the autocorrelation of the ITI time series (Figure 1D). When tapping to tones, the lag-one autocorrelation (AC-1) is typically negative, as it is here, indicating that successive ITIs are inversely correlated (i.e., a long ITI is usually followed by a shorter ITI). This result has often been interpreted as a sign of active error correction in synchronization (Repp 2005; Wing & Beek, 2002). Tapping to visual flashes (bottom row) was more variable, as has been reported previously (Chen, Repp, & Patel, 2002; Repp, 2003, Hove, et al., 2010). In particular, the RP values were more variable, as measured by a larger value of circular standard deviation of the relative phase ( $sd_{\text{circ}}$ ) and were not as closely aligned to the flash onsets (mean RP different from zero; Figure 1F). The ITI time series showed a greater variation (Figure 1G), and successive ITIs were positively correlated (AC-1  $> 0$ ; Figure 1F), indicating that the ITI was slowly drifting around the mean value and suggesting that the

error-correcting timing mechanism observed for tapping to auditory stimuli is not as effective when tapping to visual stimuli.

Differences in synchronization between tones and flashes were reliable at the group level in terms of  $sd_{\text{circ}}$  (median tone: 0.044, flash: 0.116; Wilcoxon signed rank  $Z = 126.5$ ,  $P < 0.0001$ ) and AC-1 (mean tone:  $-0.223$ , flash: 0.108;  $t(17) = 6.2$ ,  $P < 0.0001$ ).<sup>1</sup> In addition, the percentage of failures to synchronize was vastly different, with no failures to synchronize to the tones, but an average of 28% failed trials for flashes.

### 3.2 Synchronization to Moving, Colliding Visual Stimuli vs. Auditory Tones

In contrast to the poor performance for tapping to flashes, changing the nature of the visual stimulus to include movement and collisions resulted in improved synchronization performance that was substantially better than tapping to flashes and approached the performance observed when tapping to auditory stimuli. First, unlike synchronization with flashes, the circular standard deviation ( $sd_{\text{circ}}$ ) of tapping to a bouncing ball was similar to the  $sd_{\text{circ}}$  of tapping to an auditory metronome (Figure 2A). The median  $sd_{\text{circ}}$  for the bouncing ball was larger than for the tones, but this difference was not statistically significant at a population level (median 0.052 cycles [31.8 ms] vs. 0.044 cycles [26.4 ms], respectively; Wilcoxon rank sum  $P = 0.09$ ). However, given the group trend for slightly higher variability when tapping to the ball than to the tone, we also analyzed this comparison on a subject-by-subject basis. Figure 2B (bottom) shows a histogram of the within-subject difference of  $sd_{\text{circ}}$  for bouncing ball minus tones across subjects. The large modal peak centered at zero difference indicates that for the majority of subjects, performance was equivalent, or nearly so, for the auditory and visual stimuli. The long right tail shows that a minority subjects did markedly worse when tapping to the bouncing ball. These subjects ensured that the distribution was significantly biased towards better performance for tones ( $P = .011$ , Wilcoxon signed rank test;  $P = .033$  when the two outliers at the right end of the distribution are excluded). Performance for the bouncing ball is markedly superior in contrast to the same measure in flashes: Figure 2B, top shows the within-subject difference of  $sd_{\text{circ}}$  for flashes minus tones, which was always positive, and had no values near zero. Overall, this result suggests that for most individuals, the quality of timing information delivered by the bouncing ball is similar to that delivered by the discrete tones.

Tapping to the bouncing ball and tones were also equivalent in terms of the short-term autocorrelation of the inter-tap-interval (ITI) time series (Figure 2C). The mean AC-1 for both the ball and the tones was significantly negative and almost identical numerically ( $-0.221$  and  $-0.235$ , both  $P < 0.0001$  compared to zero), which is consistent with rapid serial correction of deviations from the target ITI, a pattern reported in prior tapping studies with auditory metronomes (Repp, 2005). The AC-1 values for the ball and tones were not significantly different ( $t(20) = 0.28$ ,  $P = 0.78$ ), indicating that a comparable timing process

<sup>1</sup>Unless otherwise noted, all  $t$  tests were two-tailed. Flash results were pooled across central and peripheral locations, yielding six trials to the other conditions' three. This imbalance was verified not to affect the results by reanalyzing only the central flashes. Differences remained highly significant. Non-parametric statistics were used for  $sd_{\text{circ}}$  because it departed from normality: Jarque-Bera  $P = 0.002$ .

may be involved in synchronization to the moving visual stimulus and to tones. Furthermore, individual participants' performance on the bouncing ball and tones as measured by  $sd_{circ}$  were highly correlated (Kendall's tau = 0.51,  $P < 0.001$ ), in contrast with previous findings that performance on discrete and continuous timing tasks are not correlated within individuals (Zelaznik, et al., 2002). The lack of correlation in previous work was used to support the existence of separate timing mechanisms for discrete and continuous movements. Thus, the current finding of correlation would seem to rule out the possibility that timing to a moving visual stimulus was accomplished using a continuous mode of timing.

One interesting difference emerged in the higher-order autocorrelations (lag >1), as seen in Figure 2C: the lag-3 autocorrelation (AC-3) was significantly negative for tapping with tones, but not for tapping with the bouncing ball (stars on plot; mean AC-3 for tone:  $-0.115$ ,  $t(20)=7.42$ ,  $P < 0.0001$ ; bouncing ball:  $0.0$ ,  $t(20)=0.01$ ,  $P = 0.99$ ). AC-2 was also slightly negative for tapping with tones (mean:  $-0.06$ ,  $t(20)=4.28$ ,  $P < 0.001$ ). Autocorrelations at all other higher-order lags were not significantly different from zero (all  $P > 0.54$ ). Confirming the difference in higher-order autocorrelations, a 2-way (condition x lag) repeated-measures mixed-effects model demonstrated a main effect of stimulus condition (tone vs. bouncing ball) ( $P = 0.016$ ) and a condition x lag interaction ( $P = 0.034$ ), and a marginal effect of lag ( $P=0.08$ ). A post-hoc test between conditions revealed a significant difference only for AC-3 ( $t(20)=3.53$ ,  $P < 0.002$ , but no difference between conditions for the other lags (all  $P_s > 0.33$ ). The general finding of nonzero higher order autocorrelations means that the timing of a tap depends not only on the immediately preceding interval, but that it is a more complex function of earlier taps. This pattern of findings thus suggests that the timing mechanisms involved when tapping to tones are more complex than those involved with tapping to the bouncing ball. The specific finding of a negative AC-3 replicates past experiments on the metrical organization of taps into groups of two (Vorberg & Hambuch, 1978; see section 2.4), suggesting that participants perceived the tones, but not the bouncing ball, in groups of two. This pattern will be discussed further in section 3.4, below.

### 3.3 Visuomotor Synchronization in Hearing vs. Deaf Participants

When deaf participants synchronized with isochronous flashes, the quality of their synchronization was different in a number of ways from that of hearing participants. First, deaf participants did not show the drift in their tapping that characterizes tapping by hearing individuals, as evident in autocorrelation functions (see the arrows in Figure 3A, B). For the hearing group, tapping with flashes had a positive lag-1 autocorrelation (AC-1) indicating that ITIs drifted slowly around the target value of 600 ms (mean AC-1:  $0.108$ , significantly different from zero,  $t(17)=2.7$ ;  $P = 0.024$ ). The size and significance of the positive AC-1 is consistent with past work (Chen et al., 2002) and is qualitatively different from the negative AC-1 seen when hearing individuals tap with tones (Figures 1 and 2). In contrast to the hearing participants, the mean AC-1 for deaf participants was significantly negative ( $-0.060$ ; one-sided  $t(19) = 2.27$ ,  $P = 0.017$ ) and was significantly less than the mean for the hearing group (Figure 3C; 2-sample  $t(36) = 3.56$ ;  $P = 0.001$ ). A repeated measures mixed effect model (participant group x flash position with subject as a random effect) confirmed the strong group difference when tapping to flashes ( $P = 0.002$ ). Contrary to expectations,

however, there was no difference between central and peripheral flashes ( $P = 0.55$ ), nor any interaction between group and flash position ( $P = 0.67$ ).

Deaf participants were also more often able to successfully synchronize with visual flashes. Hearing participants failed to synchronize in 29.4% of trials. Deaf participants failed about half as often, in 15.2% of trials. A 2-way binomial mixed effects linear model (group x flash position, subject as a random factor) showed a significant effect of participant group ( $P = 0.023$ ) but no trend for better performance for centrally presented flashes ( $P = 0.14$ ), and no interaction between flash position and participant group ( $P = 0.93$ ).

A third group difference we observed was in the mean asynchrony (i.e., the time interval between tap and flash), which was smaller for deaf participants ( $-7.3$  ms) than for hearing participants ( $-39.0$  ms). The asynchrony for deaf participants did not differ significantly from zero ( $t(21) = 0.54$ ,  $P = 0.60$ ), while the asynchrony for hearing participants did ( $t(21) = 3.19$ ,  $P = 0.004$ ). There was a weak main effect of participant group (2-way repeated-measures mixed-effects model, participant group:  $P = 0.065$ ). There was no effect of flash location ( $P = .40$ ), nor any interaction between group and flash location ( $P = 0.78$ ). This pattern of results suggests that deaf participants systematically tapped closer in time to the flash. The mean asynchrony for deaf participants tapping to visual flashes was not significantly different from the mean asynchrony of hearing participants tapping to an auditory metronome ( $-7.3$  ms vs.  $-18.2$  ms;  $t(42) = 0.7$ ,  $P = 0.49$ ).

In sum, on a number of measures of visual flash synchronization, the deaf participants performed significantly better than hearing participants: fewer failed trials, no ITI drift, and a trend toward smaller asynchrony. In contrast, flash location did not differentially affect the deaf participants, contrary to our prediction (both groups tended to perform better with centrally presented flashes but there was no interaction effect).

Interestingly, these differences between deaf and hearing participants in tapping to visual flashes were observed despite the fact that there was no significant group difference in the overall variance of tapping to flashes, as measured by the circular standard deviation of RP ( $sd_{\text{circ}}$ ) values ( $sd_{\text{circ}}$  median 0.116 and 0.128 for hearing and deaf, respectively; Wilcoxon rank sum  $P = 0.89$ ). That is, deaf and hearing participants showed a similar overall “spread” of temporal asynchronies between taps and flashes, but a different temporal structure of successive inter-tap intervals, and closer synchrony to flash onsets.

Deaf participants also were tested with the bouncing ball stimuli. Unlike for flashes, in which several performance measures showed advantages for deaf over hearing participants, deaf participants were no better or worse than hearing participants in tapping to the bouncing ball in terms of variability ( $sd_{\text{circ}}$ ) and AC-1 of tap time series (see Figure 4). The median  $sd_{\text{circ}}$  was 0.049 cycles [29.4 ms] for deaf participants vs. 0.044 cycles [26.4 ms] for hearing participants (Wilcoxon Rank sum  $P = 0.90$ ), and was similarly skewed, with a few outliers at the high end in each group. The mean AC-1 for deaf participants was similarly negative:  $-0.211$  vs.  $-0.235$  for hearing participants ( $t(42) = 0.34$ ,  $P = 0.73$ ). These results indicate that accurate synchronization to moving rhythmic visual stimuli does not depend

upon auditory experience and that the specialized visual experience of deaf signers does not enhance this synchronization performance.

### 3.4. Higher-order autocorrelation in synchronization with auditory and visual stimuli

Synchronization is often thought to involve only a serial error correction mechanism, in which the timing of a tap depends only on the previous interval, as typified by negative lag-1 autocorrelation of the ITI time series and zero autocorrelation at higher lags (Repp, 2005; Wing, 2002). As reported above, we found non-zero higher-order autocorrelations for hearing participants tapping to tones (Fig. 2C), specifically a negative lag-3 autocorrelation (AC-3), indicating a more complex timing control of taps that depends not only on the previous tap interval, but is influenced by ITIs farther back in time. In hearing participants, negative AC-3 was observed only for tapping to tones, but not for the visual bouncing ball. Higher-order autocorrelations have implications for the cognitive mechanisms of time keeping and have been suggested to indicate grouping or error correction at multiple metrical levels when tapping to tones (Wing & Beek, 2002). Consistent with this, negative AC-3 (with zero values for other high-order lags) has been previously observed during explicit tapping in a binary meter (Vorberg & Hambuch, 1978), and was suggested to represent motor planning not only at the tap-to-tap level, but also at the higher level of pairs of taps. Thus an interpretation of our results is that for hearing participants the metronome may have been perceived metrically, while the bouncing ball was not.

Interestingly, unlike hearing participants, deaf participants *did* show a significantly negative AC-3 when tapping with the bouncing ball (Figure 4B), which suggests that a higher-level metrical grouping/error correction may have been present in synchronization with the bouncing ball for the deaf participants (AC-3 mean  $-0.076$ ,  $t(22)=2.70$ ,  $P = 0.012$ ; AC-2, AC-4, and AC-5 not different from zero, all  $P_s > 0.51$ ). In support of this, the autocorrelation function for deaf participants synchronizing to the bouncing ball was not significantly different from that observed for hearing participants tapping to an auditory metronome. A 2-way mixed-effects model (Group x Lag, subject as random factor) revealed an effect of Lag ( $P < 0.001$ ), as expected, but demonstrate the similarity of the groups with no main effect of group ( $P = 0.17$ ), nor an interaction ( $P = 0.17$ ). Post-hoc tests confirm the theoretically important AC-3 did not differ between groups (Figure 4B, arrow,  $t(43)=1.14$ ,  $P = 0.26$ ; other lags: AC-2:  $t(43)=2.21$ ,  $P = 0.032$ ; AC-4 and AC-5, all  $P_s > 0.44$ ). This finding shows that the bouncing ball tapping performance of the deaf participants was similar to hearing participants' tapping with auditory tones, and was different from hearing participants' performance with the bouncing ball.

## 4. Discussion

The current work has shown that the long-observed auditory advantage in rhythmic synchronization is more stimulus- and experience-dependent than has been previously reported. First, we found that a silent moving visual stimulus can drive synchronization in some cases as accurately as sound. To our knowledge, this is the first demonstration that a purely visual stimulus can drive discrete rhythmic synchronization almost as accurately as an auditory metronome in the general population. Second, several measures indicated that

deaf individuals were as good as or better than hearing individuals in synchronizing movements to visual stimuli. Thus, lack of auditory experience did not impair visual synchronization ability. In contrast, it actually led to improvements in temporal processing of visual flashes, and to tapping to moving visual stimuli in a manner that closely resembled hearing participants' tapping to tones.

#### 4.1 Moving, colliding visual stimuli largely eliminate the auditory advantage in synchronization

For hearing participants, synchronization of tapping to a bouncing ball and to a tone were similar in terms of the variability of tapping, had equally high success rates for synchronization, and had the same time structure of tapping (identical lag-1 autocorrelation). Together these findings suggest that the quality of estimation of discrete points in time can be equivalent for auditory tones and moving object collisions, and that these temporal estimates are able to drive synchronized movement with similar accuracy and temporal pattern of tapping. It is possible that accurate motor synchronization to the ball occurred because the motion trajectory realistically resembled that of a moving object with sharply defined collision points and/or because relatively slower repetition rates were used compared to previous work which found an auditory advantage (e.g. Hove et al., 2010). More generally, the brain's ability to estimate time of impact for realistically-moving objects, and to coordinate motor behavior with these estimates, may be quite high due to the evolutionary significance of this ability for survival (e.g., in throwing or dodging projectiles) and because of the experience that most humans have with coordinating movement with estimated impact time of real objects (as when catching a ball). Further research is required to determine whether physically-realistic motion trajectories are the most critical factor in creating equivalent auditory and visual synchronization, and if synchronization at higher rates maintains its parity with auditory stimuli.<sup>2</sup>

Recently, research directly inspired by the present findings used a multimodal conflict paradigm to demonstrate that for expert perceivers (musicians and video-game experts), a bouncing ball stimulus was as effective at disrupting synchronization with an auditory tone as the converse (Hove, Iversen, Zhang, & Repp, 2013). Crucially, equal disruption is only predicted if the temporal precision of the representation of both stimuli is the same, in accordance with optimal integration theory, which states that the underlying temporal precision of the stimuli, rather than stimulus modality, determines the dominance of one modality over another in temporal integration tasks (Alais & Burr, 2004; Burr & Alais, 2006). This finding thus agrees with the conclusions of the present study.

The present finding of nearly equal accuracy for synchronization with moving visual and auditory targets is consistent with the hypothesis that there is a common, amodal timing mechanism that can be driven effectively by moving visual and auditory stimuli (Merchant, et al., 2013). A shared timing mechanism is consistent with the observation that participants' performance on auditory and moving visual synchronization was correlated (see Zelaznik et al., 2002), although future work is needed to determine if this is truly due to shared

---

<sup>2</sup>Note that future work with higher-rate realistically moving stimuli may be constrained by slow video refresh rates, which render fast continuous motion disjointedly with large spatial jumps, and may instead best be tested with physically moving visual stimuli.

underlying timing mechanisms. Such work could also help determine where in the brain timing information might converge across modalities. Some key differences may arise given that the visual system appears to require the prospective information provided by a moving trajectory to drive rhythmic movement with temporal accuracy, and (for hearing individuals) visual rhythms may not be perceived metrically. A recent study by Hove and colleagues (2013) suggests that one component of motor timing circuitry, the putamen, reflects the stability of synchronization, and not the modality of the stimulus, consistent with the view that amodal motor timing is driven with varying degrees of effectiveness by the different modalities.

Interestingly, when hearing participants tapped to tones, but not the bouncing ball, we observed a negative higher-order autocorrelation at lag-3 (AC-3). Why did we observe a nonzero autocorrelation at lag-3 but not at other higher-order lags? In general non-zero higher-lag autocorrelations have been proposed as signs of metrical time keeping, specifically interval timing that operates at higher metrical levels (Keller & Repp, 2004; Vorberg & Hambuch, 1978). In particular, our finding of a negative AC-3 directly replicates an earlier study of binary metrical synchronization (Vorberg & Hambuch, 1978) and thus suggests that participants may have been perceiving the auditory stimulus in a binary manner. It is well known that auditory rhythms can be perceived metrically, i.e., with periodicities at multiple time scales, as this forms the basis of much music. This pattern can occur even for simple metronomes, in which identical tones can be heard as metrical groups of strong and weak beats (Bolton, 1894), sometimes called the ‘tick-tock’ effect. The absence of higher-order autocorrelation when tapping to the bouncing ball is especially interesting in this light. Overall the findings suggest that for hearing participants tapping to the bouncing ball, while similar to the tones in the basic accuracy of synchronization, may not have yielded a metrical percept. While requiring further investigation, this in turn would suggest that for hearing individuals, a moving visual stimulus may not have equal access to neural mechanisms that are involved in more complex modes metrical timing perception (cf. Patel & Iversen, 2014).

We have focused on the discrete nature of action timing, but a finding by Hove et al. (2010) is important to note. They found that the compatibility of the direction of motion of a moving visual stimulus and the direction of motion of the tapping finger influenced synchronization success, suggesting that in some situations discrete timing may be influenced by continuous information present in the stimulus in addition to discrete timing information (Repp & Steinman, 2010). Thus, tapping to a moving visual stimulus may represent a mix of continuous and discrete timing, although our experiment does not show this. At present, it is possible to conclude that the “auditory advantage” in rhythmic synchronization is not absolute and is stimulus-dependent.

#### 4.2 Deafness enhances visuomotor synchronization

Deaf individuals synchronized more successfully with rhythmic flashes than did hearing individuals. Specifically, they had fewer outright failures to synchronize, tapped more closely in time to the flash onset, and did not display the positive lag-1 autocorrelation indicative of the drifting timing seen when hearing participants tap to flashes. Instead they

had a slightly negative lag-1 autocorrelation. For the bouncing ball, deaf individuals synchronized with equal accuracy as hearing individuals. These results indicate that the auditory scaffolding hypothesis (Conway et al, 2009) may not extend to the temporal sequential processing involved in visuomotor synchronization. Our findings clearly demonstrate that perception and synchronization with visual isochronous sequences do not require experience with auditory rhythms. It remains possible that deaf individuals may recruit auditory neural mechanisms to process visual temporal information. Several studies report that auditory cortex in congenitally deaf individuals responds to visual input (Cardin et al., 2013; Finney, Fine, & Dobkins, 2001; Karns, Dow, & Neville, 2012; Scott et al., 2014). It is possible that engagement of these “auditory” neural mechanisms may be critical for accurate visual synchronization, although without further research, this hypothesis remains speculative.

We found that auditory deprivation, perhaps in concert with life-long signing, appears to enhance some aspects of visual-motor timing behavior. This result is somewhat surprising in light of a range of findings suggesting temporal processing impairments in deaf individuals (e.g., Bolognini et al., 2012; Conway et al., 2009; Kowalska & Szelag, 2006). However, these other studies dealt with temporal sequencing or the perception of the duration of single time intervals, suggesting that the timing processes involved in synchronization with repetitive visual stimuli are not similarly impaired by the lack of auditory experience.

Enhancement of visual processing in deaf signers is generally thought to occur primarily for motion processing in the peripheral visual field and to be limited to attentional tasks (Bavelier et al., 2006). The fact that we did not observe differences in group performance based on eccentricity suggests that such specializations did not play a critical role in visuomotor synchronization to flashing stimuli. Recently, Pavani and Bottari (2012) argued that stimulus eccentricity is not a definitive predictor of whether deaf and hearing participants will differ in performance. Several studies have reported superior performance on speeded visual detection tasks when stimuli are presented either foveally or parafoveally (Bottari et al., 2010; Chen, Zhang, & Zhou, 2006; Reynolds, 1993). We suggest that the enhanced visual sensitivity observed for deaf individuals at both central and peripheral locations may underlie in part their enhanced ability to accurately tap to visual flashes. Enhanced early sensory detection (“reactivity”) may cascade through the system, facilitating higher-level visuomotor synchronization processes.

Despite finding a deaf advantage in tapping with flashes, it must be noted that the overall variability of tapping to flashes for both groups was still markedly greater than what was observed for synchronization with auditory tones in hearing participants. This suggests that despite some advantages, the experience of deaf participants was not sufficient to make tapping with flashing lights approach the accuracy observed in tapping with sound.

Further clues regarding the origin of differences due to deafness come from our observation that when tapping with the bouncing ball, deaf participants had the unusual feature of higher-order negative lag-3 autocorrelation that was also seen when hearing participants tapped with an auditory, but not visual, stimulus. Because such higher-order autocorrelations, and lag-3 in particular, have been proposed as a sign of binary hierarchical



timekeeping (as discussed in section 4.1), the finding suggests that deaf individuals might organize visual rhythm metrically, perceiving the bounces in groups of two in the way that hearing people may do for sound, but presumably without auditory imagery. This result contrasts with past suggestions that visual stimuli can only be perceived metrically if they are internally recoded into auditory images, either through obligatory mechanisms, or due to prior exposure to equivalent auditory rhythms (Glenberg & Jona, 1991; Guttman et al., 2005; Karabanov, Blom, Forsman, & Ullén, 2009; McAuley & Henry, 2010; Patel et al., 2005). Further study of naturally hierarchical visual stimuli in hearing individuals, such as musical conducting gestures (Luck & Sloboda, 2009), dance movements (Stevens, Schubert, Wang, Kroos, & Halovic, 2009), or point-light images of biological motion (Su, 2014) could reveal whether access to metrical processing advantages is more modality-independent than previously believed. Along these lines, Grahn (2012) has suggested that for hearing individuals perception of a step-wise rotating visual bar shows evidence of metricity.

Why might deaf synchronization to visual stimuli more closely resemble in quality and complexity hearing synchronization to tones? Auditory deprivation has been found to lead to enhanced visual responses in multimodal association cortices (Bavelier et al., 2001), including posterior parietal cortex, an area that is active during auditory synchronization and thought to be important for auditory to motor transformations (Pollok, Butz, Gross, & Schnitzler, 2007). A speculative possibility arising from our findings is that it is these higher-order areas that underlie the more ‘auditory-like’ character of visual synchronization in deaf individuals, perhaps because visual inputs have developed a stronger ability to drive association areas involved in beat perception. These changes may occur because of an absence of competing auditory inputs that would normally dominate multimodal areas in hearing individuals (Bavelier et al., 2006; cf. Patel & Iversen, 2014). Future neuroimaging studies could determine whether visual processing by deaf individuals engages a similar brain network as beat-based processing in hearing individuals, and if processing of moving visual stimuli shows greater commonality with auditory processing networks than does processing of flashing visual stimuli (Chen, Penhune, & Zatorre, 2008; Grahn et al., 2011; Grahn & Brett, 2009; Teki, Grube, Kumar, & Griffiths, 2011; Hove et al., 2013).

It remains possible that some of these effects on visual synchronization by deaf individuals could be driven by sign language experience, rather than by deafness. Like the current study, previous studies examining temporal processing in congenitally deaf signers did not test separate groups of hearing native signers or “oral” deaf individuals who do not know sign language (e.g., Bolognini et al., 2011; Kowalska & Szelag, 2006), possibly because these participant groups are difficult to recruit. The recent review of visual abilities in deaf individuals by Pavani and Bottari (2012) reported that the vast majority of existing studies tested congenital or early deaf participants who primarily used sign language and that differences in sign language background alone could not account for variability in study results. Enhancements that have clearly been shown to be tied to sign language experience rather than to deafness generally involve high-level visual skills, such as mental rotation (e.g., Keehner & Gathercole, 2007), visual imagery (e.g., Emmorey, Kosslyn, & Bellugi, 1993), and face processing (e.g., McCullough & Emmorey, 1997). Nonetheless, further study is needed to determine whether deafness is necessary or sufficient to bring about the observed changes in visual synchronization ability.

### 4.3 Conclusions

Our results challenge the long-held notion of fixed differences between auditory and visual rhythmic sensorimotor integration. The findings constrain cognitive models of timing by demonstrating the feasibility of an amodal timing system for visual stimuli that are appropriately structured. The fact that deaf individuals exhibited evidence of higher-level metrical processing for visual flashes provides further support for the hypothesis that an amodal timing mechanism can support rhythmic sensorimotor integration. Further, equally accurate performance by deaf and hearing participants for moving visual stimuli indicates that visual rhythms do not need to be recoded into auditory imagery, although it is nonetheless possible that deaf individuals might recruit auditory cortex when processing visual rhythms.

The fact that deafness (perhaps in conjunction with life-long signing) improved the temporal processing of visual flashes has important implications for the nature of cross-modal plasticity and for the auditory scaffolding hypothesis (Conway et al. 2009). First, the results provide evidence that cross-modal plasticity effects are not limited to attentional changes in the visual periphery and are consistent with recent research showing enhanced visual “reactivity” in deaf individuals (Bottari et al., 2010; 2011). Second, the results show that auditory deprivation during childhood does not lead to impaired temporal processing of rhythmic visual stimuli. This finding implies that the auditory scaffolding hypothesis does not apply to general timing ability and may be specifically limited to sequencing behavior.

### Acknowledgments

Supported by Neurosciences Research Foundation as part of its program on music and the brain at The Neurosciences Institute, where A.D.P. was the Esther J. Burnham Fellow, and by the National Institute for Deafness and Other Communication Disorders, R01 DC010997, awarded to K.E. and San Diego State University.

### References

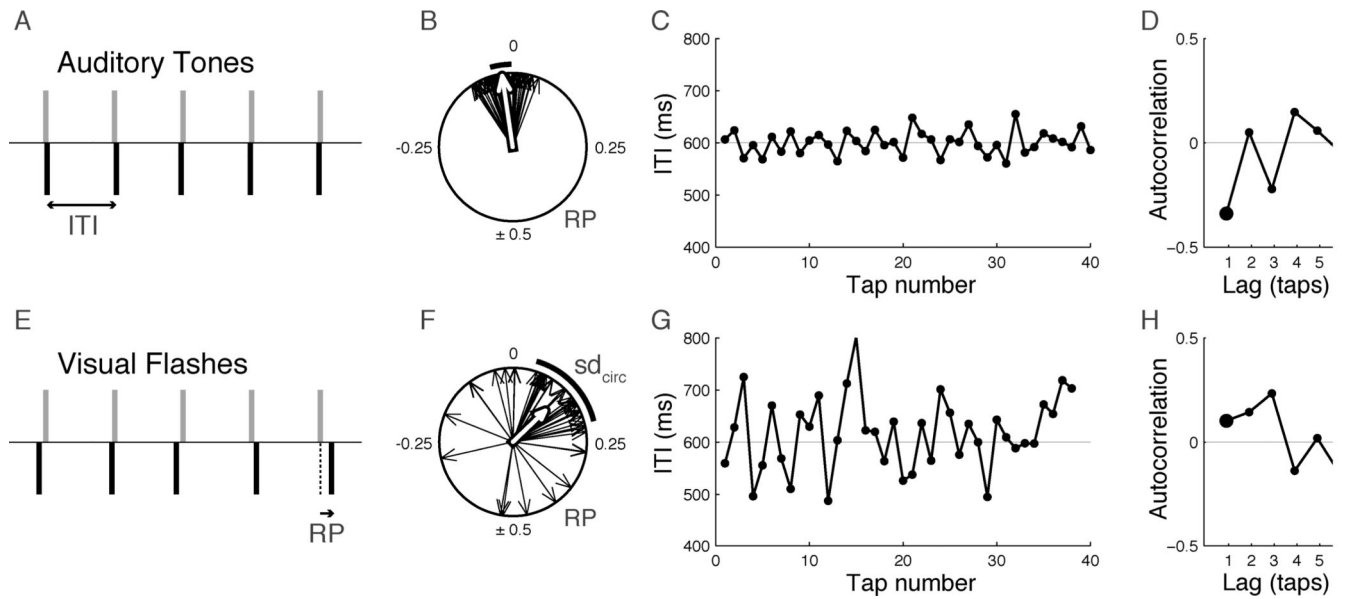
- Alais D, Burr D. The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*. 2004; 14(3):257–262. [PubMed: 14761661]
- Amazeen PG, Schmidt RC, Turvey MT. Frequency detuning of the phase entrainment dynamics of visually coupled rhythmic movements. *Biological Cybernetics*. 1995; 72(6):511–518. [PubMed: 7612722]
- Armstrong BA, Neville HJ, Hillyard SA, Mitchell TV. Auditory deprivation affects processing of motion, but not color. *Brain Research Cognitive Brain Research*. 2002; 14(3):422–434. [PubMed: 12421665]
- Bavelier D, Brozinsky C, Tomann A, Mitchell T, Neville H, Liu G. Impact of early deafness and early exposure to sign language on the cerebral organization for motion processing. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*. 2001; 21(22):8931–8942. [PubMed: 11698604]
- Bavelier D, Dye MWG, Hauser PC. Do deaf individuals see better? *Trends in Cognitive Sciences*. 2006; 10(11):512–518. [PubMed: 17015029]
- Bolognini N, Cecchetto C, Geraci C, Maravita A, Pascual-Leone A, Papagno C. Hearing shapes our perception of time: temporal discrimination of tactile stimuli in deaf people. *Journal of Cognitive Neuroscience*. 2012; 24(2):276–286. [PubMed: 21916563]
- Bolognini N, Rossetti A, Maravita A, Miniussi C. Seeing touch in the somatosensory cortex: A TMS study of the visual perception of touch. *Human Brain Mapping*. 2011; 32(12):2104–2114. [PubMed: 21305659]

- Bolton TL. Rhythm. *The American Journal of Psychology*. 1894; 6(2):145–238.
- Bootsma RJ, van Wieringen PCW. Timing an attacking forehand drive in table tennis. *Journal of Experimental Psychology: Human Perception and Psychophysics*. 1990; 16:21–29.
- Bosworth RG, Dobkins KR. Visual Field Asymmetries for Motion Processing in Deaf and Hearing Signers. *Brain and Cognition*. 2002; 49(1):170–181. [PubMed: 12027401]
- Bottari D, Caclin A, Giard M-H, Pavani F. Changes in Early Cortical Visual Processing Predict Enhanced Reactivity in Deaf Individuals. *PLoS ONE*. 2011; 6(9):e25607. [PubMed: 21980501]
- Bottari D, Nava E, Ley P, Pavani F. Enhanced reactivity to visual stimuli in deaf individuals. *Restorative Neurology and Neuroscience*. 2010; 28(2):167–179. [PubMed: 20404406]
- Burr D, Alais D. Combining visual and auditory information. *Progress in Brain Research*. 2006; 155:243–258. [PubMed: 17027392]
- Chen JL, Penhune VB, Zatorre RJ. Moving on time: Brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *Journal of Cognitive Neuroscience*. 2008; 20(2):226–239. [PubMed: 18275331]
- Cardin V, Orfanidou E, Rönnerberg J, Capek CM, Rudner M, Woll B. Dissociating cognitive and sensory neural plasticity in human superior temporal cortex. *Nature Communications*. 2013; 12(4):1473.
- Chen Y, Repp BH, Patel AD. Spectral decomposition of variability in synchronization and continuation tapping: Comparisons between auditory and visual pacing and feedback conditions. *Human Movement Science*. 2002; 21(4):515–532. [PubMed: 12450682]
- Chen Q, Zhang M, Zhou Z. Effects of spatial distribution of attention during inhibition of return on flanker interference in hearing and congenitally deaf people. *Brain Research*. 2006; 1109:117–127. [PubMed: 16859649]
- Collier GL, Logan G. Modality differences in short-term memory for rhythms. *Memory & Cognition*. 2000; 28(4):529–538. [PubMed: 10946536]
- Conway CM, Karpicke J, Anaya EM, Henning SC, Kronenberger WG, Pisoni DB. Nonverbal cognition in deaf children following cochlear implantation: Motor sequencing disturbances mediate language delays. *Developmental Neuropsychology*. 2011; 36(2):237–254. [PubMed: 21347923]
- Conway CM, Pisoni DB, Kronenberger WG. The Importance of sound for cognitive sequencing abilities: The auditory scaffolding hypothesis. *Current Directions in Psychological Science*. 2009; 18(5):275–279. [PubMed: 20725604]
- Emmorey K, Kosslyn SM, Bellugi U. Visual imagery and visual-spatial language: enhanced imagery abilities in deaf and hearing ASL signers. *Cognition*. 1993; 46(2):139–181. [PubMed: 8432094]
- Finney EM, Fine I, Dobkins K. Visual stimuli activate auditory cortex in the deaf. *Nature Neuroscience*. 2001; 4:1171–1173.
- Fisher, NI. *Statistical Analysis of Circular Data*. New York: Cambridge University Press; 1993.
- Gault RH, Goodfellow LD. An empirical comparison of audition, vision, and touch in the discrimination of temporal patterns and ability to reproduce them. *The Journal of General Psychology*. 1938; 18:41–47.
- Glenberg AM, Jona M. Temporal coding in rhythm tasks revealed by modality effects. *Memory & Cognition*. 1991; 19(5):514–522. [PubMed: 1956312]
- Grahn JA. See what I hear? Beat perception in auditory and visual rhythms. *Experimental Brain Research*. 2012; 220:51–61. [PubMed: 22623092]
- Grahn JA, Henry MJ, McAuley JD. fMRI investigation of cross-modal interactions in beat perception: audition primes vision, but not vice versa. *NeuroImage*. 2011; 54(2):1231–1243. [PubMed: 20858544]
- Grahn J, Brett M. Impairment of beat-based rhythm discrimination in Parkinson's disease. *Cortex*. 2009; 45(1):54–61. [PubMed: 19027895]
- Guttman SE, Gilroy LA, Blake R. Hearing what the eyes see: Auditory encoding of visual temporal sequences. *Psychological Science*. 2005; 16(3):228–235. [PubMed: 15733204]
- Handel S. Space is to time as vision is to audition: seductive but misleading. *Journal of Experimental Psychology: Human Perception and Performance*. 1988; 14(2):315–317. [PubMed: 2967884]

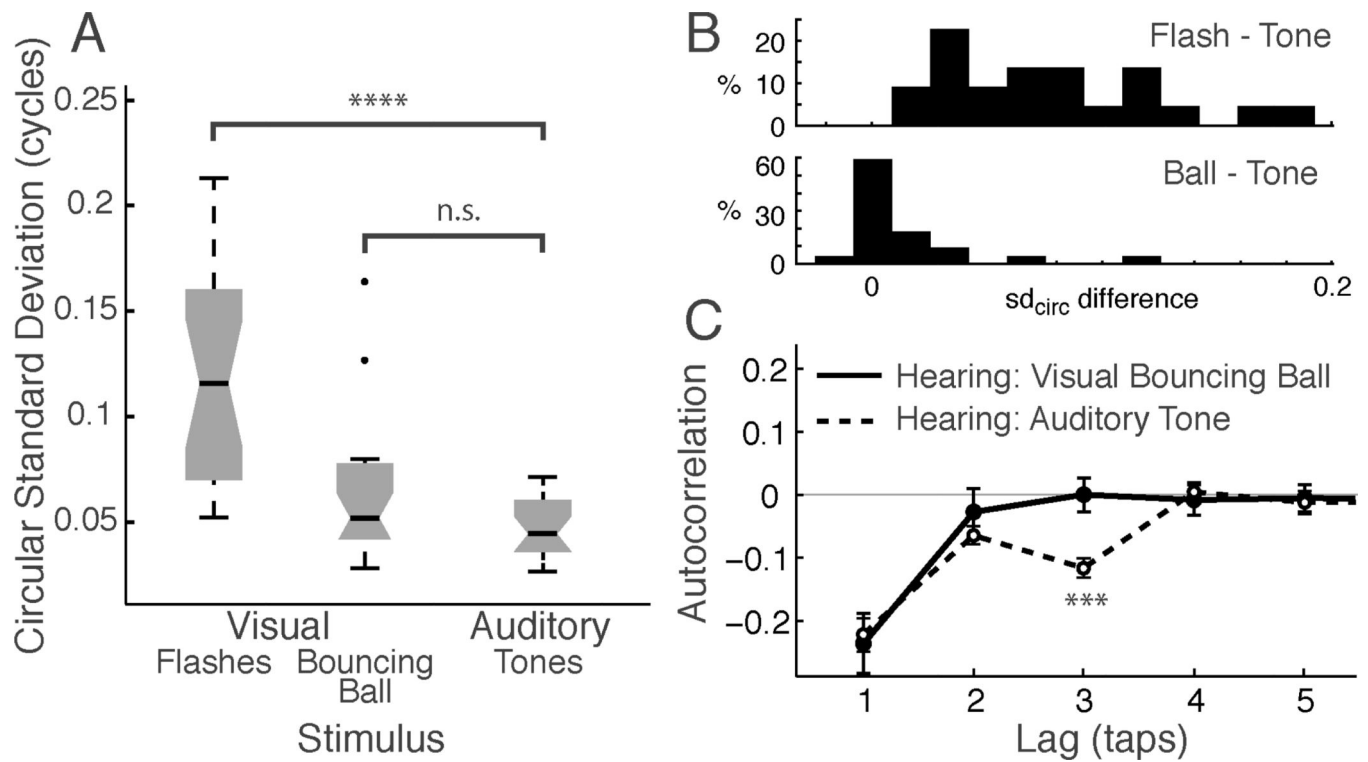
- Holcombe AO. Seeing slow and seeing fast: two limits on perception. *Trends in Cognitive Sciences*. 2009; 13(5):216–221. [PubMed: 19386535]
- Hove MJ, Fairhurst MT, Kotz SA, Keller. Synchronizing with auditory and visual rhythms: An fMRI assessment of modality differences and modality appropriateness. *NeuroImage*. 2013; 67:313–321. [PubMed: 23207574]
- Hove MJ, Iversen JR, Zhang A, Repp BH. Synchronization with competing visual and auditory rhythms: Bouncing ball meets metronome. *Psychological Research*. 2013; 77(4):388–398. [PubMed: 22638726]
- Hove MJ, Spivey MJ, Krumhansl CL. Compatibility of motion facilitates visuomotor synchronization. *Journal of Experimental Psychology: Human Perception and Performance*. 2010; 36(6):1525–1534. [PubMed: 20695698]
- Huys R, Studenka BE, Rheume NL, Zelaznik HN, Jirsa VK. Distinct timing mechanisms produce discrete and continuous movements. *PLoS Computational Biology*. 2008; 4(4):e1000061. [PubMed: 18437236]
- Jäncke L, Loose R, Lutz K, Specht K, Shah NJ. Cortical activations during paced finger-tapping applying visual and auditory pacing stimuli. *Brain Research Cognitive Brain Research*. 2000; 10(1–2):51–66. [PubMed: 10978692]
- Karabanov A, Blom Ö, Forsman L, Ullén F. The dorsal auditory pathway is involved in performance of both visual and auditory rhythms. *NeuroImage*. 2009; 44(2):480–488. [PubMed: 18848999]
- Karns CM, Dow MW, Neville HJ. Altered cross-modal processing in the primary auditory cortex of congenitally deaf adults: A visual-somatosensory fMRI study with a double-flash illusion. *Journal of Neuroscience*. 2012; 32(28):9626–9638. [PubMed: 22787048]
- Keehner M, Gathercole SE. Cognitive adaptations arising from nonnative experience of sign language in hearing adults. *Memory & Cognition*. 2007; 35(4):752–761. [PubMed: 17848032]
- Keller PE, Repp BH. Staying offbeat: Sensorimotor syncopation with structured and unstructured auditory sequences. *Psychological Research*. 2004; 69(4):292–309. [PubMed: 15616863]
- Kowalska J, Szegel E. The effect of congenital deafness on duration judgment. *Journal of Child Psychology and Psychiatry*. 2006; 47(9):946–953. [PubMed: 16930389]
- Kubovy M. Should we resist the seductiveness of the space:time::vision:audition analogy? *Journal of Experimental Psychology: Human Perception and Psychophysics*. 1988; 14(2):318–320.
- Loke WH, Song S. Central and peripheral visual processing in hearing and nonhearing individuals. *Bulletin of the Psychonomic Society*. 1991; 29:437–440.
- Luck G, Sloboda JA. Spatio-temporal cues for visually mediated synchronization. *Music Perception*. 2009; 26(5):465–473.
- McAuley JD, Henry MJ. Modality effects in rhythm processing: Auditory encoding of visual rhythms is neither obligatory nor automatic. *Attention, Perception and Psychophysics*. 2010; 72(5):1377–1389.
- McCullough S, Emmorey K. Face processing by deaf ASL signers: evidence for expertise in distinguished local features. *Journal of Deaf Studies and Deaf Education*. 1997; 2(4):212–222. [PubMed: 15579849]
- Merchant H, Harrington DL, Meck WH. Neural basis of the perception and estimation of time. *Annual Review of Neuroscience*. 2013; 36:313–336.
- Miner, JB. Motor, visual and applied rhythms (Doctoral dissertation). New York, NY: Columbia University; 1903. p. 1-121.
- Neville HJ, Lawson D. Attention to central and peripheral visual space in a movement detection task: an event-related potential and behavioral study. II. Congenitally deaf adults. *Brain Research*. 1987; 405(2):268–283. [PubMed: 3567605]
- Neville H, Bavelier D. Human brain plasticity: evidence from sensory deprivation and altered language experience. *Progress in Brain Research*. 2002; 138:177–188. [PubMed: 12432770]
- Patel AD, Iversen JR. The evolutionary neuroscience of musical beat perception: The Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in Systems Neuroscience*. 2014; 8:57. [PubMed: 24860439]

- Patel AD, Iversen JR, Chen Y, Repp BH. The influence of metricality and modality on synchronization with a beat. *Experimental Brain Research*. 2005; 163(2):226–238. [PubMed: 15654589]
- Pavani, F.; Bottari, D. Visual abilities in individuals with profound deafness: A critical review. In: Murray, MM.; Wallace, MT., editors. *The neural bases of multisensory processes*. Boca Raton, FL: CRC Press; 2012.
- Pollok B, Butz M, Gross J, Schnitzler A. Intercerebellar coupling contributes to bimanual coordination. *Journal of Cognitive Neuroscience*. 2007; 19(4):704–719. [PubMed: 17381260]
- Proksch J, Bavelier D. Changes in the spatial distribution of visual attention after early deafness. *Journal of Cognitive Neuroscience*. 2002; 14(5):687–701. [PubMed: 12167254]
- Recanzone GH. Auditory influences on visual temporal rate perception. *Journal of Neurophysiology*. 2003; 89(2):1078–1093. [PubMed: 12574482]
- Regan D. Visual judgements and misjudgements in cricket and the art of flight. *Perception*. 1992; 21:91–115. [PubMed: 1528708]
- Repp BH. 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*. 2003; 35(4):355–370. [PubMed: 14607773]
- Repp BH. Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review*. 2005; 12(6):969–992. [PubMed: 16615317]
- Repp BH, Penel A. Rhythmic movement is attracted more strongly to auditory than to visual rhythms. *Psychological Research*. 2004; 68(4):252–270. [PubMed: 12955504]
- Repp BH, Steinman SR. Simultaneous event-based and emergent timing: Synchronization, continuation, and phase correction. *Journal of Motor Behavior*. 2010; 42(2):111–126. [PubMed: 20189907]
- Repp BH, Su Y-H. Sensorimotor synchronization: A review of recent research (2006–2012). *Psychonomic Bulletin & Review*. 2013; 20:403–453. [PubMed: 23397235]
- Reynolds H. Effects of foveal stimulation on peripheral visual processing and laterality in deaf and hearing subjects. *American Journal of Psychology*. 1993; 106(4):523–540. [PubMed: 8296925]
- Scott GD, Karns CM, Stevens MW, Neville HJ. Enhanced peripheral vision processing in congenitally deaf humans is supported by multiple brain regions, including primary auditory cortex. *Frontiers in human neuroscience*. 2014; 8(177)
- Stevens C, Schubert E, Wang S, Kroos C, Halovic S. Moving with and without music: Scaling and lapsing in time in the performance of contemporary dance. *Music Perception*. 2009; 26:451–464.
- Su Y-H. Visual enhancement of auditory beat perception across auditory interference levels. *Brain and Cognition*. 2014; 90:19–31. [PubMed: 24907465]
- Teki S, Grube M, Kumar S, Griffiths TD. Distinct neural substrates of duration-based and beat-based auditory timing. *The Journal of Neuroscience*. 2011; 31(10):3805–3812. [PubMed: 21389235]
- Vorberg, D.; Hambuch, R. On the temporal control of rhythmic performance. In: Requin, J., editor. *Attention and performance VII*. Hillsdale, NJ: Erlbaum; 1978. p. 535-555.
- Wing AM. Voluntary timing and brain function: an information processing approach. *Brain and Cognition*. 2002; 48(1):7–30. [PubMed: 11812030]
- Wing AM, Beek PJ. Movement timing: A tutorial. *Attention and Performance*. 2002; 19:202–216.
- Zarco W, Merchant H, Prado L, Mendez JC. Subsecond timing in primates: Comparison of interval production between human subjects and rhesus monkeys. *Journal of Neurophysiology*. 2009; 102(6):3191–3202. [PubMed: 19812296]
- Zelaznik HN, Spencer RMC, Ivry RB. Dissociation of explicit and implicit timing in repetitive tapping and drawing movements. *Journal of Experimental Psychology: Human Perception and Performance*. 2002; 28(3):575–588. [PubMed: 12075889]

- We tested the assumed superiority of auditory over visual timing
- Colliding visual stimuli can drive synchronization nearly as accurately as sound
- Deafness leads to enhancements of visuomotor synchronization
- The results constrain theories of timing and cross-modal plasticity

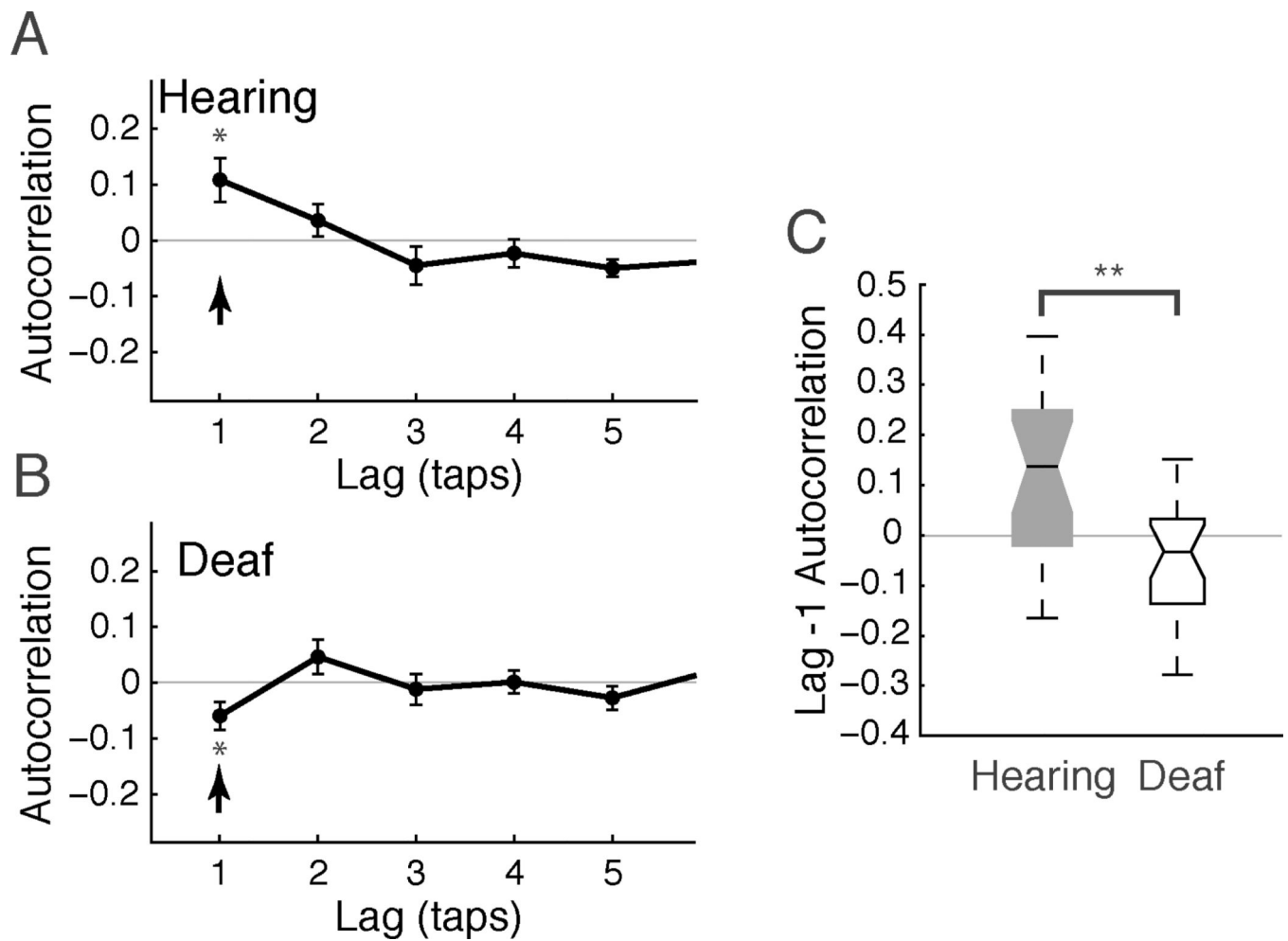


**Fig. 1.** Synchronization with visual flashes is worse than synchronization with sounds that have the same temporal characteristics. A representative example of synchronization to isochronous auditory tones (A–D) or visual flashes (E–H) is shown for a hearing participant, along with an illustration of the methods. (A, E) Schematic of the temporal relationship between stimuli (gray bars) and taps (black bars) for a few tones/flashes (ITI = inter-tap-interval). (B, F) Relative phase (RP) is plotted as vectors on a unit circle measured in cycles, where 0 indicates perfect alignment between tap and metronome, and 0.5 cycles indicates taps exactly midway between metronome events. The circular plots show examples of RP distributions for an entire trial of tapping to tones (B) and flashes (F). The white arrow indicates mean relative phase. The mean phase was negative for tones, indicating that on average, taps occurred slightly before tones. The thick arc next to each circle shows the circular standard deviation of relative phases ( $sd_{circ}$ ), a measure of the variability of relative phase/tap asynchrony. The  $sd_{circ}$  was larger for flashes. (C, G) Time series of ITIs when tapping to tones (C) or flashes (G). There was a greater degree of temporal drift when tapping to flashes. (D, H) Normalized autocorrelation of ITI time series. The lag  $-1$  autocorrelation (AC-1) is indicated by the large dot. The negative AC-1 value observed for tones (D) indicates that the durations of successive ITIs were anticorrelated, i.e., a longer ITI was followed by a shorter ITI. The positive AC-1 value seen for flashes (H) indicates that the duration of successive ITIs tended to be positively correlated, which occurs when ITI values drift slowly around the mean ITI value. Data shown are from single trials that were most representative of the group means for the hearing participants tapping to tones vs. flashes.

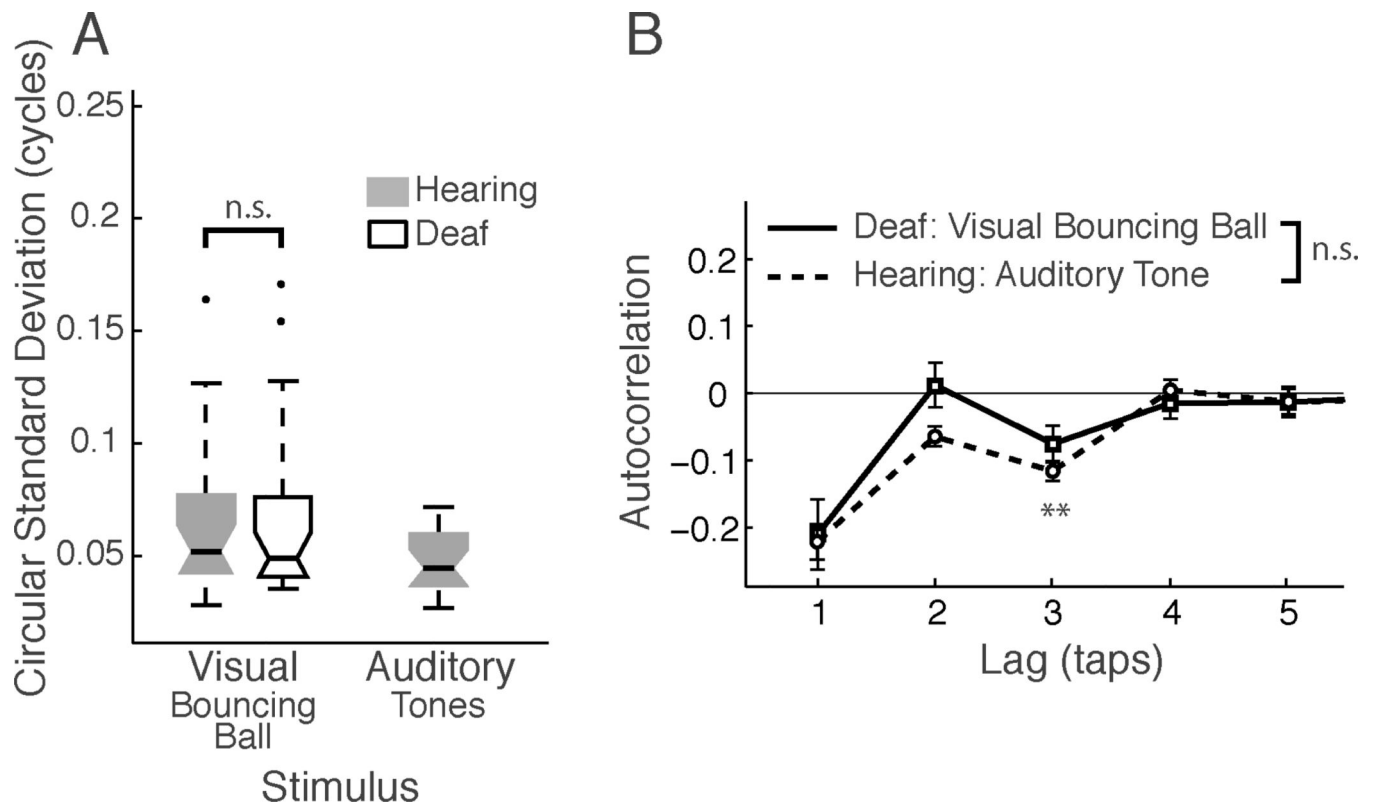


**Fig. 2.** Silent moving visual stimuli can drive synchronization nearly as accurately as sound. **(A)** Synchronization performance (circular standard deviation) for visual flashes, visual bouncing ball and auditory tones for hearing participants. Larger values indicate more variable tapping. Boxes show the median (dark horizontal line), quartiles (box; notch indicates the 95% confidence interval around the median), central 99% of data (whiskers), and outliers (dots). **(B)** Histograms of the per-subject difference in circular standard deviation ( $sd_{circ}$ ) between Flash and Tone (top) and Bouncing Ball and Tone (bottom). **(C)** Autocorrelation of ITI time series for bouncing ball (solid line, filled circles) and auditory tone (dashed line, open circles). Both conditions show marked negative Lag-1 autocorrelation, but differ in that Lag-3 autocorrelation, which was significantly negative only for tones.





**Fig. 3.** Evidence of better performance by deaf than hearing participants when tapping to flashes. **(A, B)** Group mean autocorrelation functions were characterized by mean positive lag -1 autocorrelation (AC-1) for the hearing group and mean negative AC-1 for the deaf group (arrows), indicating less drift in synchronization for the deaf participants. Error bars show the standard error of the mean. **(C)** AC-1 was significantly lower for the deaf participants. Box plot conventions as in Figure 2.



**Fig. 4.** Evidence for metricality in deaf participants' synchronization with a visual stimulus. **(A)** Tapping variability was identical for deaf and hearing participants when tapping to the bouncing ball. Box plot conventions as in Figure 2. **(B)** Autocorrelation functions for hearing participants synchronizing with tones (dashed lines) and deaf participants tapping with the bouncing ball have statistically indistinguishable negative values at lag-3 (arrow), consistent with both being perceived metrically (see text). Comparison with Figure 2B shows that this result differs from hearing participants' tapping to the bouncing ball, which does not have a negative lag-3 autocorrelation.