

Fast transfer of crossmodal time interval training

Lihan Chen · Xiaolin Zhou

Received: 19 November 2013 / Accepted: 12 February 2014 / Published online: 26 February 2014
© Springer-Verlag Berlin Heidelberg 2014

Abstract Sub-second time perception is essential for many important sensory and perceptual tasks including speech perception, motion perception, motor coordination, and crossmodal interaction. This study investigates to what extent the ability to discriminate sub-second time intervals acquired in one sensory modality can be transferred to another modality. To this end, we used perceptual classification of visual Ternus display (Ternus in Psychol Forsch 7:81–136, 1926) to implicitly measure participants' interval perception in pre- and posttests and implemented an intra- or crossmodal sub-second interval discrimination training protocol in between the tests. The Ternus display elicited either an “element motion” or a “group motion” percept, depending on the inter-stimulus interval between the two visual frames. The training protocol required participants to explicitly compare the interval length between a pair of visual, auditory, or tactile stimuli with a standard interval or to implicitly perceive the length of visual, auditory, or tactile intervals by completing a non-temporal task (discrimination of auditory pitch or tactile intensity). Results showed that after fast explicit training of interval discrimination (about 15 min), participants improved their ability to categorize the visual apparent motion in Ternus displays, although the training benefits were mild for visual timing

training. However, the benefits were absent for implicit interval training protocols. This finding suggests that the timing ability in one modality can be rapidly acquired and used to improve timing-related performance in another modality and that there may exist a central clock for sub-second temporal processing, although modality-specific perceptual properties may constrain the functioning of this clock.

Keywords Crossmodal transfer · Perceptual training · Time interval · Apparent motion · Ternus display

Introduction

Crossmodal transfer is the ability to convey information acquired in one sensory modality to another. For example, in infants, the identification of shape by touch can facilitate the gaze response to similar shapes presented in the visual modality (Gottfried et al. 1977; Hermelin and O'Connor 1964). The ability to form high-order crossmodal associations provides the basis for human cognitive development (Rosenblum et al. 2007; Rudell and Teuber 1964). Crossmodal transfer has also been explored in the temporal processing domain (Block et al. 2010; Grondin 2010; Grondin and Rousseau 1991; Ivry and Schlerf 2008; Killeen and Fetterman 1988; Lewis and Miall 2003; Rammsayer and Ulrich 2001; Zakay and Tsal 1989). The transferability of temporal processing addresses an outstanding question in sensory neuroscience: Whether the perceived timing of events is mediated by a central, supramodal timing mechanism or whether there exist multiple modality-specific systems (Buhusi and Meck 2005; Ivry and Schlerf 2008; Lewis and Miall 2003; Merchant et al. 2008; Shih et al. 2009; Zakay and Tsal 1989). A central timing mechanism

L. Chen · X. Zhou (✉)
Department of Psychology, Center for Brain and Cognitive Sciences, Peking University, Beijing 100871, China
e-mail: xz104@pku.edu.cn

L. Chen · X. Zhou
Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing 100871, China

X. Zhou
PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing 100871, China

would predict that the timing ability acquired from one sensory modality can be generalized to the processes of timing in other modalities. In contrast, the modality-specific systems or multiple clocks would allow only limited transfer of the timing ability between different sensory modalities.

Along the time scale, the range of sub-second is essential for many important sensory and perceptual tasks including speech perception, motion perception, and motor coordination. It is also essential for coordinating crossmodal interaction and multisensory integration (Alais et al. 2010; Edwards et al. 2002; Michon 1985; Schirmer 2004; Stein and Stanford 2008). There have been increasing amounts of studies investigating the underlying mechanisms of sub-second crossmodal integration or transfer. These studies employed essentially two paradigms: sensory adaptation and perceptual learning.

Crossmodal adaptation to time intervals has been implemented with the presentation of the adaptation stimuli from one modality first and the test stimuli in another modality later. The adaptation stimuli are generally paired with the test stimuli with no explicit processing requirement on the adaptation stimuli. The adaptation takes a variety of paradigms: *cross-sensory lag adaptation* (Harrar and Harris 2008; Di Luca et al. 2009), *action-event causality* (Heron et al. 2009; Sugano et al. 2010, 2012), and *temporal ventriloquist effect* (Alais and Burr 2004; Bruns and Getzmann 2008; Chen and Zhou 2011; Freeman and Driver 2008; Getzmann 2007; Kafaligonul and Stoner 2010; Shi et al. 2010; Vroomen and de Gelder 2000, 2003; see Chen and Vroomen 2013 for a review). Results of these studies, however, are largely contradictory: While some studies suggested a high-level supramodal recalibration mechanism (Heron et al. 2009), other studies showed asymmetric transfer between modalities (Harrar and Harris 2008; Sugano et al. 2010, 2012); in addition, other studies indicated a general transfer effect that is dependent on the relative temporal precision associated with the individual modality (Di Luca et al. 2009; Freeman and Driver 2008; Zhang et al. 2012). For example, Freeman and Driver (2008) repeatedly presented visual bars at two different visual fields, with fixed time lags between them. When only visual bars were presented, the equal inter-intervals between the consecutively presented bars elicited bistable visual apparent motion, without dominant perception of a specific direction (from left to right or from right to left). These bars can also be coupled with auditory tones with a variable time lag between the tone and the bar; this type of manipulation led to the perception that the bars were presented with a dominant direction (e.g., from the left to the right). After adapting to this presentation, participants were presented with the bars only. However, participants reported a biased direction for the bars, opposite to the perceived direction in the adaptation phase. Importantly, presenting only auditory

tones in the adaptation phase also led to the perception of biased direction for the subsequent visual stimuli.

Perceptual learning is a relatively permanent improvement (e.g., over weeks or months) in perceptual acuity as a result of consistent practice (Fahle and Poggio 2002). Many studies have investigated whether the ability of interval perception acquired in perceptual learning in one modality can be transferred to tasks in another modality. Three different types of training were used: (1) participants explicitly discriminated two intervals or two rhythms by deciding which interval was longer/shorter or which rhythm was faster/slower (Becker and Rasmussen 2007; Grondin and Ulrich 2011; Lapid et al. 2009; Nagarajan et al. 1998; Wright et al. 1997); (2) participants produced an interval equal to the preceding (standard) interval (Meegan et al. 2000); (3) participants made temporal order judgments for two events from either the same or different modalities (Alais and Cass 2010). In the post-training test, participants carried out essentially the same tasks but in a different modality. Results generally demonstrate the crossmodal transfer of interval timing. For example, Nagarajan et al. (1998) trained participants to discriminate the intervals between two pairs of vibratory pulses for 10–15 days, with one standard/base interval of either 75 or 125 ms and the other intervals of longer duration than the standard interval (determined by an adaptive psychophysics procedure). They found the training benefits of somatosensory interval discrimination generalized to auditory stimuli presented at comparable base intervals.

The two sets of previous studies, one using implicit timing tests and the other using explicit timing tasks, obtained support for the supramodal sub-second time interval processing mechanism. However, the two sets of studies differed on how fast the interval information could be transferred crossmodally. In the adaptation paradigm with paired adaptation test stimuli, the crossmodal transfer was very fast; in the perceptual learning paradigm, the training phase lasted several hours or several days. It is not clear whether in the latter paradigm, crossmodal transfer can take place rapidly. A fast learning mechanism, if it exists, would provide a means for overcoming inevitable multisensory (such as audiovisual) timing variation, serve to maximize the perceptual benefits of multisensory interaction, and therefore contribute to the flexibility needed for imminent perceptual decisions.

The aim of the present study was to provide evidence for the rapid transfer of the time interval processing ability, which was acquired through perceptual learning and transferred from auditory and tactile modality to vision. To this end, we conducted seven experiments in which the Ternus display was used to implicitly test participants' time perception ability before training and after training. Ternus apparent motion is a useful tool to examine implicit

temporal processing (Chen et al. 2010; Shi et al. 2010; Wang et al. 2014). A typical Ternus display consists of two visual frames. Each frame has two disks, with the second disk of the first frame and the first disk of the second frame being presented at the same location. Observers typically report two distinct percepts depending on the inter-stimulus interval (ISI) between the two frames. These are known as “element motion” (EM) and “group motion” (GM) (Fig. 1). Short ISIs usually give rise to the perception of EM. In other words, the outer dots are perceived as moving from left to right (or from right to left) while the center dot appears to remain static or flashing at the same location. In contrast, long ISIs give rise to the perception of GM. That is, the two dots are perceived moving together as a group (Chen et al. 2010; Kramer and Yantis 1997; Pantle and Petersik 1980; Pantle and Picciano 1976; Shi et al. 2010; Wang et al. 2014). In this task, the perception of EM or GM depends purely on the time interval between the two visual frames, and the variation in the percentage of GM reports for different ISIs would allow us to implicitly measure the participants’ perception of timing.

Importantly, we tested participants twice (pretest and posttest) with the Ternus task. During the actual experiments, participants completed one of the two potential tasks: (1) an *explicit time interval comparison task* in either the visual, auditory, or tactile modality for about 15 min (Experiments 1, 2, and 3); (2) an *implicit time interval perception task* in which participants discriminated the pitch frequencies or tactile intensities of auditory or tactile stimuli with different time intervals (Experiments 6 and 7). Control experiments (Experiments 4 and 5) were conducted to rule out the possibility that training benefits could be attributed to simple practice- or fatigue-based factors.

The empirical question was whether the time interval perceptual learning in a different modality affects the PSE (point of subjective equality at which the participant was equally likely to report the two percepts of apparent motion) and JND (just noticeable difference, indicating the resolution of apparent motion discrimination) for perceiving GM and EM in the Ternus display. If a particular training protocol in a modality can improve participants’ perception of time interval and this ability can be transferred to the visual modality, we should observe a reduced PSE (with a shorter ISI between the two visual frames) as well as a reduced JND for the perception of GM in the Ternus posttest relative to the pretest. That is, the two visual frames in the Ternus display should be more easily “separated,” and the sensitivity of discrimination should be elevated in the posttest as compared with the pretest. Moreover, if the training benefit is dependent upon the efficacy of the interval training protocol, we should observe larger transfer benefits for explicit timing training protocols (Experiments 1, 2, and 3) than for the implicit timing training protocols

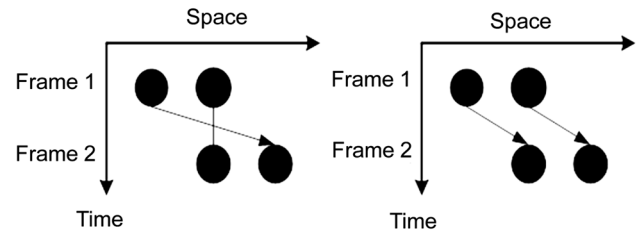


Fig. 1 The Ternus display. Two possible motion percepts: **a** Element motion (EM) for short ISIs with the middle disk being perceived as static and the outer disk being perceived as moving from one side to the other. **b** Group motion (GM) for long ISIs with the two disks being perceived moving together as a group

(Experiments 6 and 7); this is because in the former protocols, timing was task relevant and the focus of attention.

Methods

Participants

To avoid potential contamination of tasks, we adopted a between-participants design for training protocols. Eighty-five undergraduate and graduate students (23 females, aged between 20 and 34 years, average age 24.5 years) participated in the seven experiments. Twelve participants were tested for Experiment 1 (auditory interval training), 12 for Experiment 2 (tactile interval training), 13 for Experiment 3 (visual interval comparison), 12 for Experiment 4 (Ternus task—three repetitions), 13 for Experiment 5 (Ternus task—two repetitions), 11 for Experiment 6 (auditory pitch discrimination), and 12 for Experiment 7 (tactile intensity discrimination).

All the participants reported normal hearing and somatosensory perception and were naïve to the purpose of the study. The experiment was performed in compliance with the institutional guidelines set by the Ethics Committee, Department of Psychology at Peking University.

Pre- and posttests

The Ternus display was used to measure the implicit perception of time interval for pre- and posttests. Each Ternus display was composed of two frames, with each frame of two black disks (12.71 cd/m^2 in luminance) presented horizontally for 30 ms on a gray background (16.11 cd/m^2 in luminance). The two frames shared one disk location at the center of the screen and contained the other two disks on the horizontally opposite side of the center (Ternus 1926; Shi et al. 2010). The diameter of each black disk was 1.6° in visual angle, and the distance between the centers of the two adjacent disks was 3.1° . The duration of each frame was 30 ms.

For a given trial, the ISI between the two visual frames of the Ternus display was randomly selected from one of the following seven durations: 50, 80, 110, 140, 170, 200, or 230 ms. There were 40 trials for each level of ISI, with 20 trials for the leftward or rightward directions of apparent motion. The presentation order of the 280 trials was randomized for each participant. These trials were divided into 4 blocks, with 70 trials for each block. Participants took short breaks between blocks.

Training protocols

For crossmodal explicit interval training, the interim training phase task was to judge which of the two intervals was longer, with each of the intervals marked either by two sound beeps (Experiment 1) or by two tactile taps (Experiment 2). A standard interval of 140 ms between the two beeps or between two tactile taps was included in each trial, and the comparison interval varied from 50 to 230 ms, with 30 ms as the step size. The presentation order of the standard interval and the comparison intervals was randomized. Each auditory beep was 65 dB, lasting 30 ms, and its frequency was 1,000 Hz. The beeps were conveyed through binaural headphones. The tactile stimuli were produced using solenoid actuators with embedded cylinder metal tips taped to the tip of the left middle finger (Heijo Box, Heijo Research Electronics, UK). The maximum contact area was about 4 mm² and the maximum output is 3.06 W. The intensity of tactile tap was calibrated for each participant and was set at a medium intensity level before the formal experiment.

For intra-modal training, the task was to judge which of the two visual intervals was longer, with each interval marked by two consecutive visual dots at the fixation position (Experiment 3). A standard interval of 140 ms between the two visual dots was included in each trial, and the comparison interval varied from 50 to 230 ms, with 30 ms as the step size.

Experiments 4 and 5 were conducted as controls. Different from the above training protocols, the interim training phase required participants to either perform the Ternus task, as in the pre- and posttests, or simply to rest.

For the training protocol of Experiment 6, two pairs of pure tones were presented in each trial, with one pair having a standard pitch frequency (1,000 Hz) and another pair having a frequency varying from 850 to 1,150 Hz, with 50 Hz as the step size. Participants were asked to decide which tone pair had a higher frequency. For the training protocols of Experiment 7, two pairs of tactile stimulation were administered as taps to the tip of the left middle finger, with one pair of taps of standard intensity and another pair of variable intensities. The tactile intensity was varied from Level 1 (the weakest) to Level 7 (the strongest stimulation).

Level 4 was the standard intensity. Importantly, in both protocols, the standard pair of stimuli had a fixed time interval of 140 ms, while the comparison stimuli had a variable interval selected randomly from a range between 50 and 230 ms. Each stimulus lasted 30 ms. Therefore, the presentation of the training stimuli was essentially the same as in Experiments 1 and 2, but the time intervals between the paired stimuli were task irrelevant.

Training in each experiment (except Experiment 5) was composed of 168 trials, divided into 6 blocks. Participants could take a break between blocks. The whole training session for an experiment lasted about 15 min. The rest time between pre- and posttest in Experiment 5 was also about 15 min.

Procedure

Prior to the formal experiment, participants practiced to get familiar with the Ternus displays of either typical “EM” (with ISI = 50 ms) or typical “GM” (ISI = 260 ms). They were asked to discriminate the two types of motions by pressing the left or right mouse button. The mapping between response button and the type of motion was counterbalanced across participants. During practice, an incorrect response was resulted in immediate feedback on the screen revealing the correct response (EM or GM). The practice session continued until the participant’s accuracy of report was close to 100 %. All the participants could achieve an accuracy of 95 % within 120 trials.

For the Ternus motion discrimination task in all the experiments, each trial began with a fixation cross-presented at the center of the screen for 300 ms. A blank display (with a gray background) was then shown for a random duration of 300–500 ms, followed by a Ternus display with a variable ISI (50, 80, 110, 140, 170, 200, or 230 ms) between the two frames. After a blank screen of 300 ms, a question mark appeared to prompt participants to make a two-alternative forced-choice response, indicating whether they had perceived EM or GM. The next trial began 500 ms after the participant pressed the button.

For auditory interval training (Experiment 1), a trial began with the presentation of a fixation cross at the center of the screen for a random duration of 300–500 ms. A blank display (with a gray background) was then shown for 300–500 ms. Then, the participant was presented with the standard auditory time interval (i.e., the interval between two sound beeps was 140 ms) and a comparison interval (marked by two sound beeps with a random interval between 50 and 230 ms). The order of the two intervals was randomly determined, and the time interval between the sound beeps was 800–1,000 ms and was randomly determined. After another blank screen of 500–700 ms, a text appeared on the screen, asking “Which time interval is

longer, the first or the second?” The next trial began 500–700 ms after button press.

For tactile interval training (Experiment 2) and visual interval training (Experiment 3), as well as the task-irrelevant training of tone pitch discrimination (Experiment 6) and tactile intensity discrimination (Experiment 7), the procedures and the timing parameters were essentially the same as the auditory interval training in Experiment 1.

The test room was dimly lit with an average ambient luminance of 0.12 cd/m^2 . Visual stimuli were presented on a 22-in. CRT monitor ($1,024 \times 768$ pixels; 100 Hz) positioned at eye level. Viewing distance was set to 57 cm, maintained with a chin rest. A headset (Philips, SHM 1900) was used to emit sound stimuli as well as to prevent the faint noise produced by tactile taps in Experiment 2. Stimulus presentation and data collection were implemented with computer programs developed with MATLAB 7.1 (MathWorks Inc., Natick, MA) and Psychophysics Toolbox (Brainard 1997; Pelli 1997).

Results

Pretests

For each level of ISI between the two frames, the percentage of “GM” reports was collapsed over two motion directions. The seven data points (one for each ISI) were fitted into the psychometric curve using a logistic function (Treutwein and Strasburger 1999). The transitional ISI (PSE) was calculated by estimating the 50 % of reporting GE on the fitted curve. The JND was calculated as half of the difference between the lower (25 %) and upper (75 %) bound of the thresholds in the psychometric curve.

Comparisons were conducted for the PSEs derived for the seven groups of participants. There were no significant differences between PSEs in the seven experiments, $F(6,78) < 1$. The PSEs were 138.8 ± 3.9 ms for Experiment 1, 142.5 ± 4.7 ms for Experiment 2, 131.9 ± 5.9 ms for Experiment 3, 136.5 ± 6.1 ms for Experiment 4, 132.5 ± 4.1 ms for Experiment 5, 137.2 ± 3.4 ms for Experiment 6, and 129.4 ± 6.5 ms for Experiment 7.

Comparisons were also made for the JNDs, which measured the task difficulty and participants’ sensitivity in discriminating the two possible percepts of visual Ternus display. Again, there were no differences between the seven experiments, $F(6,78) < 1$. The JNDs were 36.2 ± 4.9 ms for Experiment 1, 30.4 ± 4.1 ms for Experiment 2, 32.2 ± 4.6 ms for Experiment 3, and 29.7 ± 4.1 ms for Experiment 4, 24.5 ± 2.5 ms for Experiment 5, 26.4 ± 3.1 ms for Experiment 6, and 31.5 ± 3.7 ms for Experiment 7. These null results suggested that the participants in the seven experiments were generally well matched

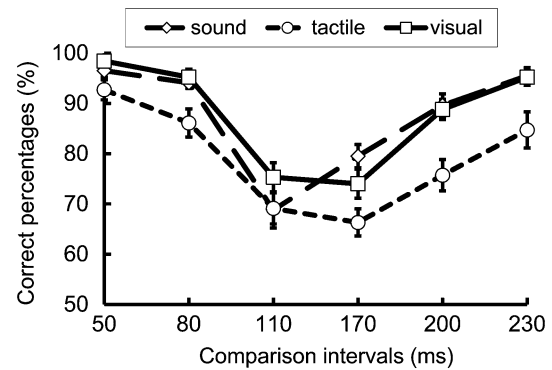


Fig. 2 The correct response percentages of interval discrimination as a function of the comparison intervals (50, 80, 110, 170, 200, and 230 ms) versus standard interval (140 ms). The dashed line with diamonds depicts the correct percentages in sound interval discrimination (Experiment 1), the dotted line with circles depicts the correct percentages in tactile interval discrimination (Experiment 2), and the solid line with squares depicts visual interval discrimination (Experiment 3)

in their ability to discriminate visual apparent motion and to implicitly process time intervals between visual frames.

Interval training

Figure 2 depicts participants’ performance on the interval training tasks in Experiments 1–3. Clearly, accuracy of responses to the interval question varied as a function of the difference between the comparison interval and the standard interval (140 ms). Repeated measures ANOVA with the comparison interval (seven levels) as a within-participant factor showed a significant main effect of interval, $F(6,66) = 292.74$, $p < 0.001$ for Experiment 1, $F(6,66) = 94.68$, $p < 0.001$ for Experiment 2, and $F(6,72) = 282.99$, $p < 0.001$ for Experiment 3.

Similar observations also apply to Experiments 6 and 7. There was a significant main effect of pitch frequency, $F(6,60) = 46.88$, $p < 0.001$, and a significant main effect of tactile intensity level, $F(6,60) = 60.44$, $p < 0.001$. It is clear from Fig. 3 that performance was worse when the pitch frequency or intensity level was close to that of the standard stimuli.

Pretests vs. posttests

Figure 4 illustrates the mean psychometric estimates (with standard errors) across participants in each of the seven experiments. Figure 5 illustrates the PSEs and JNDs for the pre- and posttests of Ternus motion in each experiment.

We first conducted a mixed-model ANOVA for Experiments 1–5, with the difference between pretests and posttests as a within-participant factor and the experiment as a between-participant factor. For PSEs, there were no

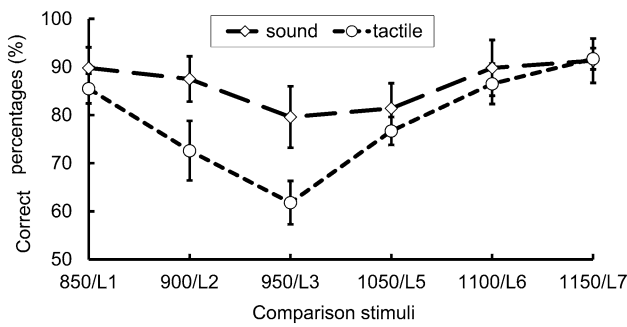


Fig. 3 The correct response percentages for pitch/intensity discriminations as a function of the pitch frequency (850, 900, 950, 1,050, 1,100, and 1,150 Hz) versus standard pitch (1,000 Hz) or as a function of tactile intensity levels (from L1 to L7 except L4) versus standard intensity (level 4). The *dashed line with diamonds* depicts the correct percentages in pitch discrimination (Experiment 6), and the *dotted line with circles* depicts the correct percentages in tactile intensity discrimination (Experiment 7)

significant main effect of test, $F(1,57) = 2.05, p > 0.1$, and no main effect of experiment, $F(4,57) < 1$. However, the interaction between test and experiment was significant, $F(4,57) = 2.72, p < 0.05$, indicating that the difference between the pre- and posttests had different patterns

in different experiments. For the JNDs, there was a significant main effect of test, $F(1,57) = 8.51, p < 0.01$, but no main effect of experiment, $F(4,57) < 1$. Importantly, the interaction between test and experiment was marginally significant, $F(4,57) = 2.28, p = 0.07$. ANOVAs including all the seven experiments obtained essentially the same patterns of results. We therefore test the effects separately for each experiment.

Experiment 1

The mean PSEs were 138.8 ± 3.9 and 129.6 ± 5.3 ms for the pretest and posttest, respectively. The difference between them was significant, $t(11) = 3.58, p < 0.01$. The difference between the mean JNDs for the pretest (36.2 ± 4.9 ms) and posttest (30.5 ± 4.2 ms) was also significant, $t(11) = 2.86, p < 0.05$. Thus, auditory interval training seemed to improve the participants' ability to discriminate the apparent motion in the Ternus display.

Experiment 2

The mean PSEs were 142.5 ± 4.7 and 132.8 ± 5.5 ms and the mean JNDs were 30.4 ± 4.1 and 25.8 ± 3.6 ms for

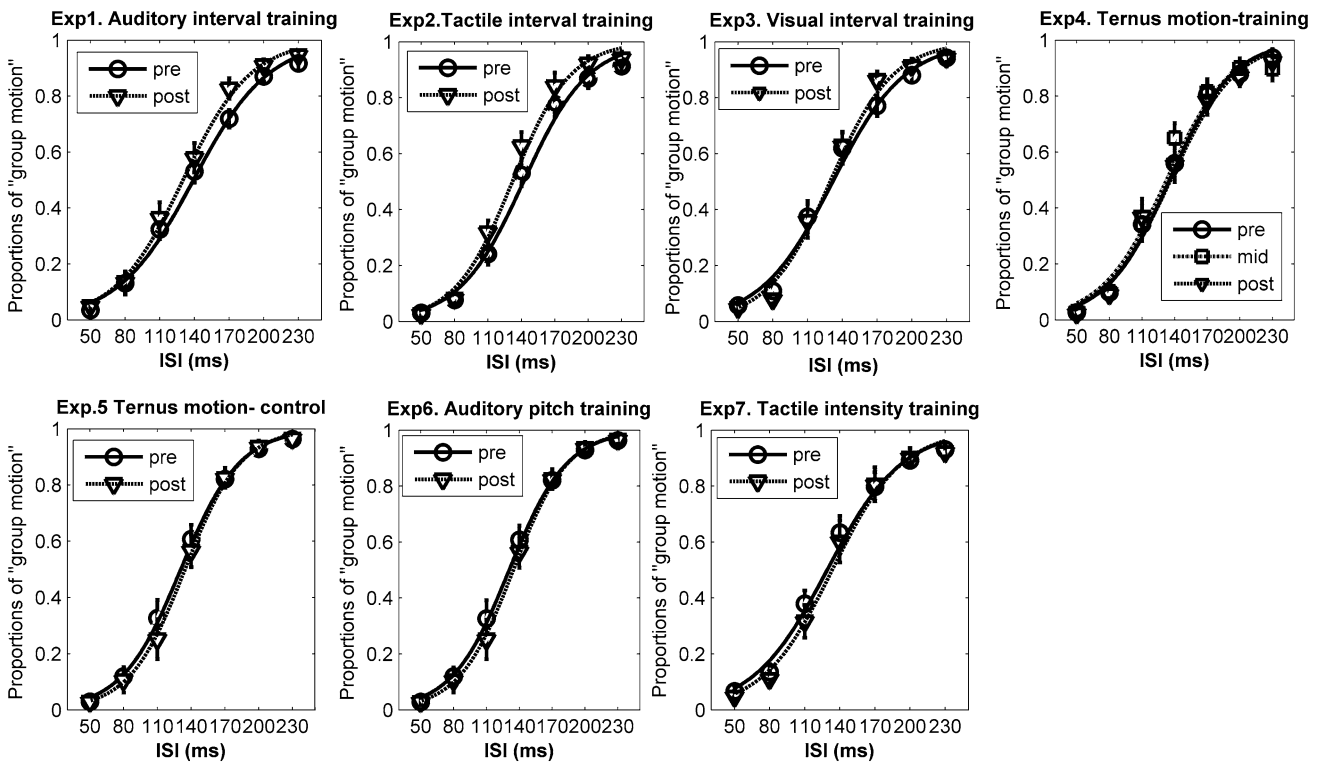
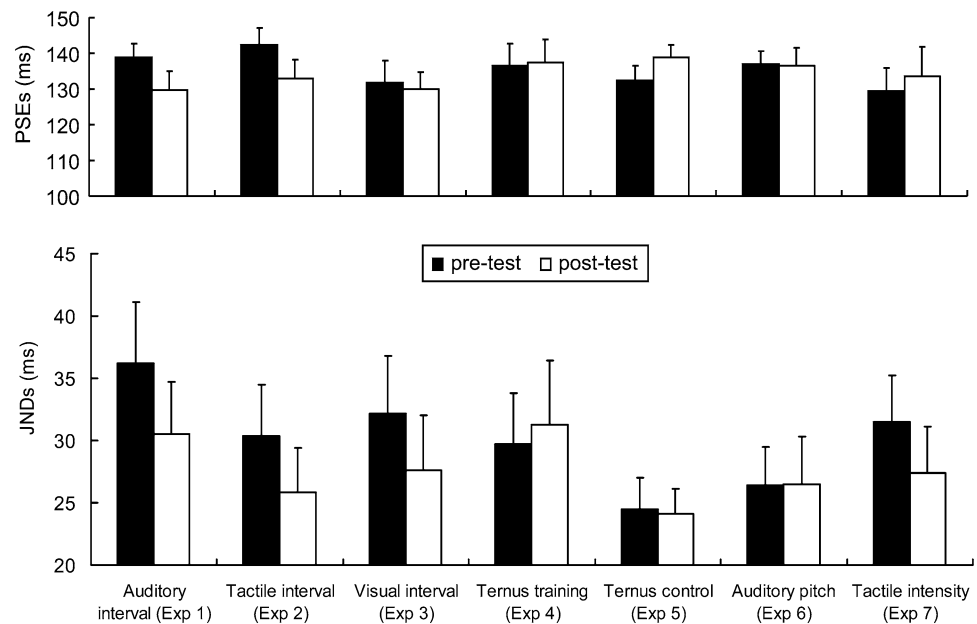


Fig. 4 Psychometric curves for mean percentages for reports of group motion (GM) in the seven experiments (pre- and posttest of the Ternus apparent motion). The *solid curve (circles)* represents the proportion of GM response as a function of the ISI between the two Ternus frames in the pretest, and the *dotted curve (with triangles)* represents GM responses in the posttest. The *error bars* represent standard errors (In Experiment 4, the Ternus motion tasks were repeated in the three sessions)

Fig. 5 PSEs and JNDs for pre- and posttests of discriminating the Ternus apparent motion in the seven experiments. The *black bars* indicate values of PSEs or JNDs in the pretests, and the *white bars* indicate values of PSEs and JNDs in the posttests. The *error bars* represent the standard errors



the pretest and the posttest, respectively. *T* tests showed significant differences between the pre- and posttests, $t(11) = 2.78$, $p < 0.05$ for PSE and $t(11) = 2.42$, $p < 0.05$ for JND. Again, tactile interval training appeared to have improved the ability to discriminate visual apparent motion for the Ternus display.

Experiment 3

The PSEs showed no significant difference between the pretest (131.8 ± 5.9 ms) and the posttest (129.9 ± 4.7 ms), $t(12) < 1$, although the JNDs did have an effect (32.2 ± 4.6 vs. 27.6 ± 4.4 ms), $t(12) = 2.55$, $p < 0.05$. Thus, unlike auditory and tactile interval training, the intra-modal training had only partial effects on the perception of visual apparent motion.

Experiment 4

We obtained the PSEs and JNDs for the three sessions of Ternus apparent motion discrimination. The PSEs were 136.5 ± 6.1 , 134.4 ± 6.7 , and 137.5 ± 6.4 ms for the pre-, interim-, and posttest, $F(2,22) < 1$. The JNDs were 29.7 ± 4.1 , 31.2 ± 6.3 , and 31.3 ± 5.1 ms, respectively, for the three sessions, $F(2,22) < 1$. Clearly, repeatedly performing the apparent motion discrimination for the Ternus display did not lead to performance improvement.

Experiment 5

The PSEs were 132.4 ± 4.1 and 138.9 ± 3.5 ms for the pretest and posttest, respectively, with a slight increase in

PSE in the post-test, $t(12) = -1.95$, $p = 0.08$; the JNDs were 24.5 ± 2.5 and 24.1 ± 2.0 ms for the pre- and posttest, $t(12) < 1$. Therefore, Experiment 5 replicated the null effect of simple repetition from Experiment 4.

Experiment 6

The PSEs showed no significant difference between the pretest (137.2 ± 3.4 ms) and the posttest (136.6 ± 4.9 ms), $t(10) < 1$; the JNDs also showed no differences between pre- and posttest (26.4 ± 3.1 vs. 26.5 ± 3.8 ms), $t(10) < 1$.

Experiment 7

There was no significant difference between the pre- and posttests in either the PSE (129.4 ± 6.5 vs. 133.6 ± 8.1 ms), $t(11) = -1.28$, $p > 0.1$, or the JND (31.5 ± 3.7 vs. 27.4 ± 3.7 ms), $t(11) = -1.81$, $p = 0.10$.

Discussion

The empirical question for this study was whether perceptual discrimination of visual apparent motion in the Ternus display could be influenced by fast training of time interval discrimination in the same or different modalities. Given that the discrimination of “EM” versus. “GM” depends upon the implicitly perceived time interval between the two visual frames of the Ternus display, any changes in sensitivity (corresponding to JND) and threshold (corresponding to PSE) in perceptual discrimination across the pre- and posttests could be attributed to the impact of crossmodal/

intra-modal interval training in between the tests. Results showed that fast, explicit training (about 15 min) on cross-modal (auditory or tactile) temporal interval discrimination improved the performance on the posttest, with significantly diminished PSEs and JNDs in the posttests (Experiments 1 and 2). The training benefit was limited for intra-modal (visual) interval training (Experiment 3), where the JNDs decreased (i.e., with enhanced sensitivity) but the PSEs were unaffected. This enhancement of sensitivity was most likely due to the increased precision in visual temporal processing that translates into the implicit interval discrimination in the visual Ternus task (Stevenson et al. 2013). Repeatedly performing the Ternus motion discrimination task, however, did not change the PSEs or JNDs across sessions (Experiments 4 and 5), thus ruling out any simple practice- or fatigue-based accounts for the transfer effects in Experiments 1–3. When timing between auditory and tactile stimulation was task irrelevant, but was varied in the same way as in the explicit training protocols (Experiments 6 and 7), no implicit learning occurred and no improvement was observed on the performance in the subsequent Ternus task.

For temporal processing across different sensory modalities, two distinct theoretical frameworks—dedicated and intrinsic models—have recently been summarized for human timing (see Ivry and Schlerf 2008). Dedicated models assume the existence of a single, centralized timing mechanism (i.e., a “central clock”). Any training benefit of temporal processing in one modality can be automatically transferred or generalized to the processing in other modalities. In contrast, the intrinsic models assume that time is an inherent property of task-related neural activity in a particular modality (Allman et al. 2014; Ivry and Schlerf 2008). Accordingly, training on temporal ability in one modality should not transfer to other modalities. In the current study, the transfer of the auditory/tactile explicit interval training to the visual events gives support for a common supramodal timing process, which is consistent with most of the previous studies reviewed in Introduction (Meegan et al. 2000; Nagarajan et al. 1998; Wright et al. 1997; see also Alais and Cass 2010 for partial transfer). However, in contrast to the lengthy training of hours or days (e.g., Stevenson et al. 2013), the benefits of the current trainings were achieved rapidly, suggesting that the ability of timing is more flexible than the most literature that has demonstrated (Alais and Cass 2010; Grondin and Ulrich 2011; Meegan et al. 2000; Nagarajan et al. 1998; Wright et al. 1997).

However, for the exposure to task-irrelevant interval variations, there were no transfer benefits in discriminating visual apparent motion, suggesting that, at least for the given short training protocols, exposure to implicit timing cannot recalibrate the sensory processing in another modality. The contrasting findings for different training protocols (explicit vs. implicit) demonstrate that the crossmodal

transfer effect in sub-second timing is task dependent and is an attention-guided process (Ahissar and Hochstein 1993, 1997, 2004; Huang et al. 2007; Huyck and Johnsrude 2012; Li et al. 2004).

The effects of fast interval training reflect the malleability of multisensory temporal processes. Crossmodal interval training changed the precision in perceiving the temporal window or time interval between the two events enclosing the time interval. Previous modeling of perceptual learning in the spatial task of orientation discrimination has indicated that the tuning curve of trained cortical neurons either amplifies or sharpens the firing rates of target neurons responsible for encoding the orientation discrimination (Bejjanki et al. 2011). By analogy, we suggest that interval training may elicit both types of changes that lead to improved performance in subsequent probe tasks of temporal processing. In this way, the explicitly acquired timing ability was subsequently used to better separate the two visual frames in the Ternus display, leading to more reports of GM. Indeed, there is evidence showing that there are partially overlapping neural circuits (particularly in the parietal cortex) for explicit timing and implicit timing across different sensory modalities (Coull and Nobre 2008; Wiener et al. 2010). It is possible that the firing of neurons after training for explicit timing is more precise when these neurons are also used for tasks involving implicit timing.

For the explicit interval discrimination protocols, participants were required to make perceptual decision to the two pairs of consecutively presented stimuli (intervals), both of which should be retained in memory. The discrimination of intervals is intrinsically task demanding, while discrimination of temporal order and perceptual decision of implicit timing may require less attentional resource and hence are less task demanding. This explains why there was no significant transfer effect for the temporal order judgment task (Alais and Cass 2010) and no effect for the implicit exposure to timing properties in Experiments 6 and 7. The training efficacy may obey the inverse effectiveness law. This law states that multisensory integration is more likely to occur when the constituent sensory stimuli presented in isolation evoke relatively weak responses (Stein and Stanford 2008). We suggest that the training benefits are inversely proportional to the task demand for unisensory interval training. When the training task is highly demanding, the transfer effect/benefit is observed (Di Luca et al. 2009; Stein and Stanford 2008). Still, one might wonder why there was no robust transfer of temporal processing for intra-modality (the visual modality) training. After all, the training protocol for the visual modality was structured in the same way as training protocol for the auditory and tactile modality and had essentially the same level of task requirements. We argue that the failure to achieve full benefit of intra-modal training was probably due to the

relatively short training protocol, which would not allow the participant to acquire the full ability to discriminate time intervals in the visual modality. Previous studies have shown that visual interval timing, when compared with auditory and tactile modalities, is associated with larger internal noise, making its estimation less accurate (Burr et al. 2009; Glenberg et al. 1989; Glenberg and Jona 1991; Grondin and Rousseau 1991; Grondin et al. 2005; Recanzone 2003; Welch and Warren 1980). It is possible that within a 15-min training session for the visual modality, the “noisy” nature has not been ameliorated and the system only partially benefits from the training protocol. In other words, although there may exist a central clock which functions universally in different modalities and which subserves the transfer of training effects across modalities, the way this central system functions in a particular modality may be partially modulated or constrained by the sensory and perceptual properties of this modality.

Note that in the current study, vision was the target modality for testing the transfer effect in all the experiments. Earlier studies reported crossmodal transfer effects from the somatosensory to the auditory modality (Nagaranjan et al. 1998) and from the auditory modality to motor timing (Meegan et al. 2000). But attempts to demonstrate crossmodal transfer from the auditory modality to vision were generally unsuccessful (Grondin and Ulrich 2011; Lapid et al. 2009). Only in a recent perceptual learning study did researchers observe the transfer of temporal discrimination training from the auditory modality to the visual modality (but not vice versa; Bratzke et al. 2012). It is for further research whether a fast transfer can be observed for other target modalities with explicit time interval training but with implicit testing of the timing ability in the target modality.

In conclusion, with fast crossmodal interval training, we provided evidence that the implicit processing of sub-second visual interval could be improved after explicit training with crossmodal interval information. We also showed that there may exist a central clock for sub-second temporal processing, although modality-specific perceptual properties may constrain the functioning of this clock.

Acknowledgments This study was supported by grants from the Natural Science Foundation of China (31200760, 30110972, 91232708) and by grants from the National High Technology Research and Development Program (863 Program: 2012AA011602) and National Basic Research Program (973 Program: 2010CB833904) of China.

References

Ahissar M, Hochstein S (1993) Attentional control of early perceptual learning. *Proc Natl Acad Sci USA* 90:5718–5722

- Ahissar M, Hochstein S (1997) Task difficulty and the specificity of perceptual learning. *Nature* 387:401–406
- Ahissar M, Hochstein S (2004) The reverse hierarchy theory of visual perceptual learning. *Trends Cogn Sci* 8:457–464
- Alais D, Burr D (2004) The ventriloquist effect results from near optimal bimodal integration. *Curr Biol* 14:257–262
- Alais D, Cass J (2010) Multisensory perceptual learning of temporal order: audiovisual learning transfers to vision but not audition. *PLoS ONE* 5(6):e11283
- Alais D, Newell FN, Mamassian P (2010) Multisensory processing in review: from physiology to behaviour. *Seeing Perceiving* 23:3–38
- Allman MJ, Teki S, Griffiths TD, Meck WH (2014) Properties of the internal clock: first- and second-order principles of subjective time. *Annu Rev Psychol* 65:743–771
- Becker MW, Rasmussen IP (2007) The rhythm aftereffect: support for time sensitive neurons with broad overlapping tuning curves. *Brain Cogn* 64:274–281
- Bejjanki VR, Beck JM, Lu ZL, Pouget A (2011) Perceptual learning as improved probabilistic inference in early sensory areas. *Nat Neurosci* 14:642–648
- Block RA, Hancock PA, Zakay D (2010) How cognitive load affects duration judgments: a meta-analytic review. *Acta Psychol* 134:330–343
- Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10:433–436
- Bratzke D, Seifried T, Ulrich R (2012) Perceptual learning in temporal discrimination: asymmetric cross-modal transfer from audition to vision. *Exp Brain Res* 221:205–210
- Bruns P, Getzmann S (2008) Audiovisual influences on the perception of visual apparent motion: exploring the effect of a single sound. *Acta Psychol* 129:273–283
- Buhusi CV, Meck WH (2005) What makes us tick? Functional and neural mechanisms of interval timing. *Nat Rev Neurosci* 6:755–765
- Burr D, Banks MS, Morrone MC (2009) Auditory dominance over vision in the perception of interval duration. *Exp Brain Res* 198:49–57
- Chen L, Vroomen J (2013) Intersensory binding across space and time: a tutorial review. *Atten Percept Psychophys* 75:790–811
- Chen L, Zhou X (2011) Capture of intermodal visual/tactile apparent motion by moving and static sounds. *Seeing Perceiving* 24:369–389
- Chen L, Shi Z, Müller HJ (2010) Influences of intra- and crossmodal grouping on visual and tactile Ternus apparent motion. *Brain Res* 1354:152–162
- Coull JT, Nobre AC (2008) Dissociating explicit timing from temporal expectation with fMRI. *Curr Opin Neurobiol* 18:137–144
- Di Luca M, Machulla T, Ernst MO (2009) Recalibration of multisensory simultaneity: cross-modal transfer coincides with a change in perceptual latency. *J Vis* 9(12):116
- Edwards CJ, Alder TB, Rose GJ (2002) Auditory midbrain neurons that count. *Nat Neurosci* 5:934–936
- Fahle M, Poggio T (2002) *Perceptual learning*. MIT Press, Cambridge, MA
- Freeman E, Driver J (2008) Direction of visual apparent motion driven by timing of a static sound. *Curr Biol* 18:1262–1266
- Getzmann S (2007) The effect of brief auditory stimuli on visual apparent motion. *Perception* 36:1089–1103
- Glenberg A, Jona M (1991) Temporal coding in rhythm tasks revealed by modality effects. *Mem Cogn* 19:514–522
- Glenberg A, Mann S, Altman L, Forman T, Prochise S (1989) Modality effects in the coding reproduction of rhythms. *Mem Cogn* 17:373–383
- Gottfried AW, Rose SA, Bridger WH (1977) Cross-modal transfer in human infants. *Child Dev* 48:18–23
- Grondin S (2010) Timing and time perception: a review of recent behavioral and neuroscience findings and theoretical directions. *Atten Percept Psychophys* 72:561–582

- Grondin S, Rousseau R (1991) Judging the relative duration of multimodal short empty time intervals. *Percept Psychophys* 49:245–256
- Grondin S, Ulrich R (2011) Duration discrimination performance: no cross-modal transfer from audition to vision even after massive perceptual learning. In: Vatakis A, Esposito A, Giagkou M, Cummins F, Papadelis G (eds) *Multidisciplinary aspects of time and time perception*, vol 6789. Lecture notes in computer science Springer, Berlin, pp 92–100
- Grondin S, Roussel ME, Gamache PL, Roy M, Ouellet B (2005) The structure of sensory events and the accuracy of time judgments. *Perception* 34:45–58
- Harrar V, Harris LR (2008) The effect of exposure to asynchronous audio, visual, and tactile stimulus combinations on the perception of simultaneity. *Exp Brain Res* 186:517–524
- Hermelin B, O'Connor N (1964) Crossmodal transfer in normal, subnormal and autistic children. *Neuropsychologia* 2:229–235
- Heron J, Hanson JVM, Whitaker D (2009) Effect before cause: supramodal recalibration of sensorimotor timing. *PLoS ONE* 4(11):e7681
- Huang X, Lu H, Tjan BS, Zhou Y, Liu Z (2007) Motion perceptual learning: when only task-relevant information is learned. *J Vis* 7:1–10
- Huyck JJ, Johnsrude IS (2012) Rapid perceptual learning of noise-coded speech requires attention. *J Acoust Soc Am* 131:236–242
- Ivry RB, Schlerf JE (2008) Dedicated and intrinsic models of time perception. *Trends Cogn Sci* 12:273–280
- Kafaligonul H, Stoner GR (2010) Auditory modulation of visual apparent motion with short spatial and temporal intervals. *J Vis* 10:1–13
- Kanai R, Lloyd H, Buetti D, Walsh V (2011) Modality-independent role of the primary auditory cortex in time estimation. *Exp Brain Res* 209:465–471
- Killeen PB, Fetterman JG (1988) A behavioral theory of timing. *Psychol Rev* 95:274–295
- Kramer P, Yantis S (1997) Perceptual grouping in space and time: evidence from the Ternus display. *Percept Psychophys* 59:87–99
- Lapid E, Ulrich R, Rammsayer T (2009) Perceptual learning in auditory temporal discrimination: no evidence for a cross-modal transfer to the visual modality. *Psychon Bull Rev* 16:382–389
- Lewis AP, Miall RC (2003) Distinct systems for automatic and cognitively controlled time measurement: evidence from neuroimaging. *Curr Opin Neurobiol* 13:250–255
- Li W, Piëch V, Gilbert CD (2004) Perceptual learning and top-down influences in primary visual cortex. *Nat Neurosci* 7:651–657
- Meegan DV, Aslin RN, Jacobs RA (2000) Motor timing learned without motor training. *Nat Neurosci* 3:860–862
- Merchant H, Zarco W, Prado L (2008) Do we have a common mechanism for measuring time in the hundreds of millisecond range? Evidence from multiple-interval timing tasks. *J Neurophysiol* 99:939–949
- Michon JA (1985) The compleat time experimenter. In: Michon JA, Jackson JL (eds) *Time, mind and behavior*. Springer, New York, pp 20–52
- Nagarajan SS, Blake DT, Wright BA, Byl N, Merzenich MM (1998) Practice-related improvements in Somatosensory interval discrimination are temporal specific but generalize across skin location, hemisphere, and modality. *J Neurosci* 18:1559–1570
- Pantle AJ, Petersik JT (1980) Effects of spatial parameters on the perceptual organization of a bistable motion display. *Percept Psychophys* 27:307–312
- Pantle AJ, Picciano L (1976) A multistable movement display: evidence for two separate motion systems in human vision. *Science* 193:500–502
- Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* 10:437–442
- Rammsayer T, Ulrich R (2001) Counting models of temporal discrimination. *Psychon Bull Rev* 8:270–277
- Recanzone GH (2003) Auditory influences on visual temporal rate perception. *J Neurophysiol* 89:1078–1093
- Rosenblum LD, Miller RM, Sanchez K (2007) Lip-read me now, hear me better later: cross-modal transfer of talker-familiarity effects. *Psychol Sci* 18:392–396
- Rudell RG, Teuber HL (1964) Crossmodal transfer of shape discrimination by children. *Neuropsychologia* 2:1–8
- Schirmer A (2004) Timing speech: a review of lesion and neuroimaging findings. *Cogn Brain Res* 21:269–287
- Shi Z, Chen L, Müller HJ (2010) Auditory temporal modulation of the visual Ternus effect: the influence of time interval. *Exp Brain Res* 203:723–735
- Shih LY, Kuo WJ, Yeh TC, Tzeng OJ, Hsieh JC (2009) Common neural mechanisms for explicit timing in the sub-second range. *NeuroReport* 20:897–901
- Stein BE, Stanford TR (2008) Multisensory integration: current issues from the perspective of the single neuron. *Nat Rev Neurosci* 9:255–266
- Stevenson RA, Wilson MM, Powers AR, Wallace MT (2013) The effects of visual training on multisensory temporal processing. *Exp Brain Res* 225:479–489
- Sugano Y, Keetels M, Vroomen J (2010) Adaptation to motor-visual and motor-auditory temporal lags transfer across modalities. *Exp Brain Res* 201:393–399
- Sugano Y, Keetels M, Vroomen J (2012) The build-up and transfer of sensorimotor temporal recalibration measured via a synchronization task. *Front Psychol* 3:1–12
- Ternus J (1926) Experimentelle untersuchungen über phänomenale identität. *Psychol Forsch* 7:81–136
- Treutwein B, Strasburger H (1999) Fitting the psychometric function. *Percept Psychophys* 61:87–106
- Vroomen J, de Gelder B (2000) Sound enhances visual perception: cross-modal effects of auditory organization on vision. *J Exp Psychol Hum Percept Perform* 26:1583–1590
- Vroomen J, de Gelder B (2003) Visual motion influences the contingent auditory motion after effect. *Psychol Sci* 14:357–361
- Wang Q, Bao M, Chen L (2014) The role of spatio-temporal and spectral cues in segregating short sound events: evidence from auditory Ternus display. *Exp Brain Res* 232:273–282
- Welch RB, Warren DH (1980) Immediate perceptual response to intersensory discrepancy. *Psychol Bull* 88:638–667
- Wiener M, Turkeltaub PE, Coslett HB (2010) Implicit timing activates the left inferior parietal cortex. *Neuropsychologia* 48:3967–3971
- Wright BA, Buonomano DV, Mahncke HW, Merzenich MM (1997) Learning and generalization of auditory temporal-interval discrimination in humans. *J Neurosci* 17:3956–3963
- Zakay D, Tsal Y (1989) Awareness of attention allocation and time estimation accuracy. *Bull Psychon Soc* 27:209–210
- Zhang H, Chen L, Zhou X (2012) Adaptation to visual or auditory time intervals modulates the perception of visual apparent motion. *Front Integr Neurosci* 6:100