

Perceptual learning in temporal discrimination: asymmetric cross-modal transfer from audition to vision

Daniel Bratzke · Tanja Seifried · Rolf Ulrich

Received: 10 April 2012 / Accepted: 21 June 2012 / Published online: 7 July 2012
© Springer-Verlag 2012

Abstract This study assessed possible cross-modal transfer effects of training in a temporal discrimination task from vision to audition as well as from audition to vision. We employed a pretest–training–post-test design including a control group that performed only the pretest and the post-test. Trained participants showed better discrimination performance with their trained interval than the control group. This training effect transferred to the other modality only for those participants who had been trained with auditory stimuli. The present study thus demonstrates for the first time that training on temporal discrimination within the auditory modality can transfer to the visual modality but not vice versa. This finding represents a novel illustration of auditory dominance in temporal processing and is consistent with the notion that time is primarily encoded in the auditory system.

Keywords Perceptual learning · Temporal discrimination · Auditory dominance

Introduction

An outstanding question in time perception research is whether training people on temporal discrimination tasks in one sensory modality transfers to other sensory modalities. Providing an answer to this question is of theoretical importance because it helps discriminate between two distinct theoretical frameworks of human timing—that is,

dedicated and intrinsic models (see Ivry and Schlerf 2008). Dedicated models of time perception assume the existence of either a single centralized timing mechanism, as, for example, in the prominent scalar expectancy theory (Gibbon et al. 1984; Wearden et al. 1998) which is rooted in classic pacemaker-counter models (e.g. Creelman 1962; Treisman et al. 1990), or multiple clocks each dedicated to a specific interval (e.g. Buhusi and Meck 2009). According to these models, the postulated timing mechanisms can operate on different sensory inputs—that is, the same neural circuit is involved in temporal discrimination, even when it receives information from different modalities (e.g. Ulrich et al. 2006). According to dedicated models, training effects thus should generalize across different modalities. In contrast, intrinsic models consider time an inherent property of neural dynamics, as, for example, in the state-dependent network model (Karmarkar and Buonomano 2007). These models imply that timing is modality-specific (see Ivry and Schlerf 2008) and thus predict that training on temporal discrimination in one modality should not transfer to other modalities.

A few studies have provided evidence for generalization of perceptual learning across modalities. In an early study, Warm et al. (1975) reported symmetric cross-modal transfer effects of training on temporal discrimination of supra-second intervals (6–18 s) from audition to vision as well as from vision to audition. In the sub-second range, cross-modal transfer effects have been reported from the somatosensory to the auditory modality (Nagarajan et al. 1998) and from the auditory modality to motor timing (Meegan et al. 2000). In this time range, however, attempts to demonstrate cross-modal transfer from auditory training to visual stimuli have yet been unsuccessful (Grondin and Ulrich 2011; Lapid et al. 2008).

D. Bratzke (✉) · T. Seifried · R. Ulrich
Cognition and Perception, Department of Psychology,
University of Tübingen, Schleichstrasse 4,
72076 Tübingen, Germany
e-mail: daniel.bratzke@uni-tuebingen.de

Concerning the interval specificity of perceptual learning, most studies using intervals in the sub-second range could not find generalization across intervals (Banai et al. 2010; Karmarkar and Buonomano 2003; Meegan et al. 2000; Nagarajan et al. 1998; Wright et al. 1997). One exception is the study of Lapid et al. (2009), who observed a transfer of training effects with an auditory 100-ms interval to an auditory 500-ms interval.

In summary, most studies on perceptual learning in temporal discrimination conducted so far revealed interval-specific cross-modal transfer of perceptual learning in temporal discrimination of sub-second intervals. Previous studies, however, failed to demonstrate such transfer effects from the auditory to the visual modality (Grondin and Ulrich 2011; Lapid et al. 2009). This appears somewhat surprising since temporal discrimination is usually much better for auditory stimuli than for visual stimuli (e.g. Grondin 1993; Penney et al. 2000; Wearden et al. 1998) and auditory training of temporal discrimination can improve motor timing (Meegan et al. 2000). Furthermore, results of Guttman et al. (2005) and Kanai et al. (2011) suggest that time information is primarily encoded in the auditory system, and for temporal discrimination, visual information is automatically transformed into an auditory representation. One would therefore rather expect cross-modal transfer effects of temporal discrimination training from the auditory to other modalities than vice versa (see also Grondin and Ulrich 2011).

The studies of Grondin and Ulrich (2011) and Lapid et al. (2009) both appear to have provided less than optimal conditions to reveal cross-modal transfer of perceptual learning. In the study of Grondin and Ulrich, participants conducted two sessions (pre- and post-test) of visual interval discrimination. Between the test sessions, one group of participants received a massive training of 1,800 trials of auditory interval discrimination in a single 2.5-h session, whereas a control group waited for the same period. Visual discrimination performance improved in both groups from the pretest to the post-test session. Although the training group showed an improvement twice as large as the one for the control group, there was no significant difference between the two groups. The massive training might have caused fatigue, which in turn might have masked potential transfer effects (see also Grondin and Ulrich 2011).

Lapid et al. (2009) employed 5 training sessions distributed across several days with a total of 3,000 trials of training. Nevertheless, they found no evidence for a transfer from training with an auditory 100-ms interval to a visual 100-ms interval. In this study, participants conducted two pretest sessions distributed across 2 days. That is, after the two pretest sessions, the control group had already received 400 trials of training in each test

condition. Possibly, this amount of training during the two pretest sessions was already sufficient to induce perceptual learning for the visual 100-ms test condition. Furthermore, Lapid et al. as well as Grondin and Ulrich (2011) did not provide performance feedback during the training sessions. Although it is widely believed that perceptual learning does not require feedback (e.g. Fahle and Edelman 1993; Shiu and Pashler 1992), there is some evidence that perceptual learning occurs faster with than without feedback (Fahle and Edelman 1993).

As we have discussed, several factors may have prevented Grondin and Ulrich (2011) and Lapid et al. (2009) from observing transfer of perceptual learning in auditory temporal discrimination to the visual modality. To overcome these limitations, the present study (a) employed only one pretest session, (b) distributed the training sessions across several days, (c) provided performance feedback during the training sessions, and (d) rigorously controlled for possible order effects in the pre- and post-test sessions.

Method

Participants

Seventy-two volunteers participated in the study. Participants reported normal hearing and normal or corrected-to-normal vision. They were randomly assigned to one of three groups (visual training, auditory training, no training). Seven participants of the original sample were replaced by new participants because their performance deviated from mean performance in at least one of the test conditions in the pretest or post-test session by more than 3 *SD*. Finally, the visual training group consisted of 5 males and 19 females (mean age, 25.5 years; *SD* = 5.8), the auditory training group of 6 males and 18 females (mean age, 24.8 years; *SD* = 4.5), and the no-training group of 3 males and 21 females (mean age, 24.1 years; *SD* = 6.4).

Apparatus and Stimuli

Participants were tested individually in a sound-attenuated, dimly illuminated booth. They sat at a distance of approximately 50 cm from a circular red LED with a diameter of 0.57°, wore headphones, and responded with two external response keys. The experiment was written in Matlab (The MathWorks 2007), using the Psychophysics Toolbox (Brainard 1997; Pelli 1997) version 3.0.8, and was run on an Apple iMac with OS X. As in previous research on transfer of temporal learning (e.g. Grondin and Ulrich 2011; Lapid et al. 2009; Nagarajan et al. 1998; Wright et al. 1997), empty intervals were used. Auditory intervals were

marked by 15-ms beeps (1 kHz) with 5-ms sinusoidal onset and offset ramps and were presented binaurally via headphones at an intensity of 70 dB SPL. The visual intervals were marked by two 15-ms light pulses produced by the LED. The length of the interval was defined as the time interval from the offset of the first marker to the onset of the second marker. The intervals were similar to the ones used in a study by Grondin and Rammsayer (2003).

Procedure

A single stimulus protocol was used (see, e.g. Karmarkar and Buonomano 2003). Each pre- and post-test as well as the training sessions started with the presentation of one of four possible standard intervals. The standard intervals were auditory intervals of either 100 ms (A100) or 200 ms (A200) and visual intervals of either 100 ms (V100) or 200 ms (V200). In each trial, participants were then presented with one of 8 possible comparison intervals that were symmetrically distributed around the respective standard interval (A100: 79, 85, 91, 97, 103, 109, 115, 121 ms; A200: 158, 170, 182, 194, 206, 218, 230, 242 ms; V100: 65, 75, 85, 95, 105, 115, 125, 135 ms; V200: 130, 150, 170, 190, 210, 230, 250, 270 ms). A larger range of comparison intervals was chosen for the visual than for the auditory modality because discrimination performance is generally better in the auditory than in the visual modality (e.g. Penney et al. 2000; Wearden et al. 1998). Participants were instructed to respond as accurately as possible with their left (right) index finger when the comparison interval was shorter (longer) than the initially presented standard interval. Subsequent trials were separated by one of four possible foreperiods (800, 900, 1,000, or 1,100 ms). Feedback was only provided during training sessions. In case of an erroneous response, participants received feedback by providing a tactile stimulation of their right lower leg for 500 ms.

In each test session (pre- and post-test), participants were tested in each of the four conditions (A100, A200, V100, and V200). The order of the conditions within test sessions was counterbalanced across participants and groups according to a complete permutation of the four conditions, that is, in each group all of the 24 possible orders of test conditions were tested. Each test session (pre- and post-test) consisted of 4 blocks (one for each testing condition) of 160 trials. Training sessions also consisted of 4 blocks of 160 trials each, but in these sessions, participants were tested only in their respective training condition. Each session lasted about 30 min. For all groups, the pretest and a post-test session took place on two consecutive Fridays. The two training groups additionally participated in 4 training sessions from Monday to Thursday between the two test sessions. The auditory training group

was trained with the A100 condition and the visual training group with the V100 condition. The two training groups received a total of 2,560 trials of training distributed across 4 days.

Data analysis

For each factorial combination of participant, test condition, and daily session, a psychometric function was traced, plotting the eight comparison intervals on the x -axis and the probability of responding “long” on the y -axis. A logistic function was fitted to the resulting curves in order to compute the maximum likelihood estimates of the standard measure of discrimination performance, that is, the difference limen (DL). The DL is estimated as being half the interquartile range of this fitted function—that is, $DL = (x_{.75} - x_{.25})/2$, where $x_{.25}$ and $x_{.75}$ denote the values of the comparison that yield 25 and 75 % “longer” responses, respectively. The Weber fraction (DL/standard duration) was then computed for each factorial combination of participant, test condition, and daily session. Finally, in order to test for differences in perceptual learning between groups and conditions, we calculated difference scores between the Weber fractions obtained in the pretest and post-test sessions.

Results

Figure 1 depicts mean Weber fractions as a function of session for the two training groups and the no-training group. For the two training groups, performance in their trained condition mainly improved from the pretest session to session 2 (i.e. the first training session) and remained rather unchanged from session 3 until the post-test session. Nevertheless, an ANOVA including only the two training groups and the training sessions (sessions 2–5) revealed a significant effect of session on the Weber fraction, $F(3, 138) = 4.14, p = .008$.

Figure 2 shows the difference scores of the Weber fractions (pretest minus post-test) as a function of test condition and group. An ANOVA with the factors group, modality, and duration for the difference scores revealed significant main effects of group, $F(2, 69) = 3.29, p = .043$, modality, $F(1, 69) = 4.95, p = .029$, and duration, $F(1, 69) = 15.60, p < .001$. As can be seen in Fig. 2, the difference scores were larger for the two training groups (visual training, 0.05; auditory training, 0.06) than for the no-training group (0.01), indicating that perceptual learning occurred in the two training groups. Difference scores were also larger for the visual stimuli than for the auditory stimuli (0.05 vs. 0.03), and they were also larger for the 100-ms intervals than for the untrained 200-ms

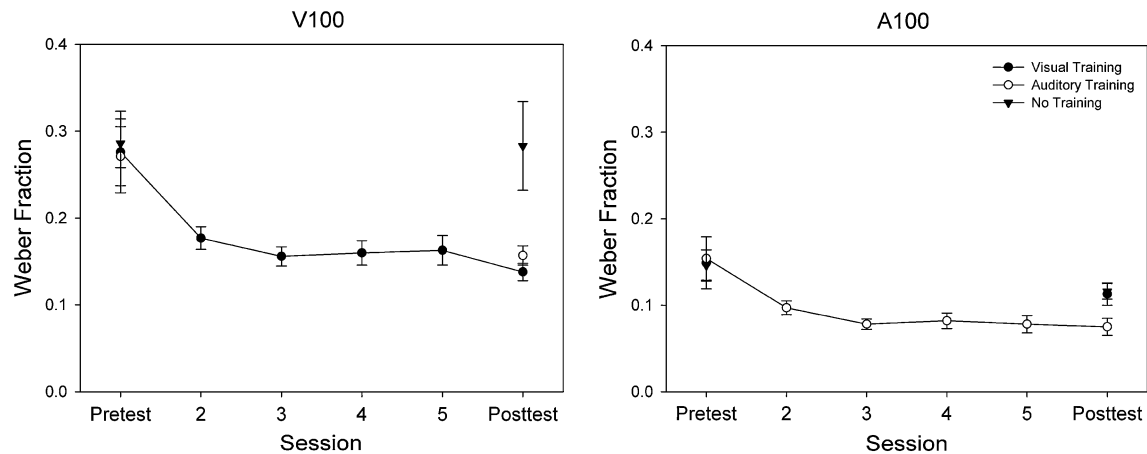


Fig. 1 Weber fractions for the visual 100-ms interval (V100) and the auditory 100-ms interval (A100) as a function of session and training group. Error bars represent ± 1 SE

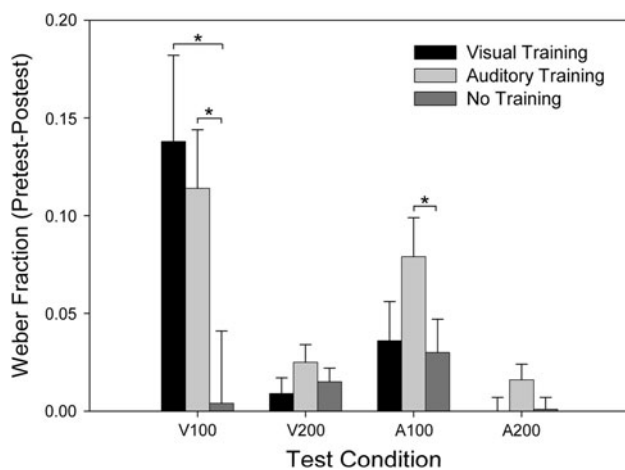


Fig. 2 Difference scores of the Weber fractions in the pretest and post-test sessions (pretest minus post-test) as a function of test condition and training group. Asterisks indicate significantly larger difference scores for the training group than for the no-training group ($p < .05$). Error bars represent ± 1 SE

intervals (0.07 vs. 0.01). Both two-way interactions including the factor group were marginally significant, group \times modality: $F(2, 69) = 2.79$, $p = .069$; group \times duration: $F(2, 69) = 2.80$, $p = .068$. Most importantly, there was a significant three-way interaction effect of group, modality, and duration on the difference scores, $F(2, 69) = 4.17$, $p = .020$. In order to explore this interaction effect in more detail, we calculated t tests (one-tailed) between each training group and the no-training group separately for each test condition. For the V100 condition, both the visual and the auditory training group showed significantly larger difference scores than the no-training group (visual training vs. no training, $t(45) = 2.32$, $p = .012$; auditory training vs. no training, $t(44) = 2.31$, $p = .013$). For the V200 condition, both comparisons were not significant (visual training vs. no training, $t(46) = 0.52$,

$p = .698$; auditory training vs. no training, $t(44) = 0.88$, $p = .192$). For the A100 condition, only the auditory training group showed larger difference scores than the no-training group (visual training vs. no training, $t(45) = 0.24$, $p = .406$; auditory training vs. no training, $t(45) = 1.87$, $p = .034$). Finally, for the A200 condition there was only a marginally significant difference between the auditory training group and the no-training group (auditory training vs. no training, $t(42) = 1.59$, $p = .059$; visual training vs. no training, $t(44) = 0.14$, $p = .555$). Taken together, perceptual learning occurred in both training groups. A cross-modal transfer of perceptual learning, however, was observed only for auditory training. Additionally, there was only little evidence for an intramodal transfer of perceptual learning to intervals of other durations.

Discussion

The present study examined cross-modal transfer effects of training on a temporal discrimination task with sub-second intervals from vision to audition as well as from audition to vision. In order to avoid limitations of previous studies (Grondin and Ulrich 2011; Lapid et al. 2009), we employed only one pretest session, distributed the training sessions across several days, provided performance feedback during the training sessions, and rigorously controlled for possible order effects in the pre- and post-test sessions.

In line with the majority of previous studies, we found that training improved discrimination performance for the trained interval (e.g. Lapid et al. 2009; Wright et al. 1997; but see Rammsayer 1994). Both the visual and the auditory training group showed larger improvement in temporal discrimination for their respective training condition than the no-training group. The main question of the present study was whether these training effects would generalize

across modalities. The results revealed a clear cross-modal transfer effect which, however, was asymmetric: auditory training transferred to the visual domain, but visual training did not transfer to the auditory domain. This finding contrasts with the results of previous studies (Grondin and Ulrich 2011; Lapid et al. 2009) that perceptual learning in temporal discrimination of auditory stimuli does not generalize to visual stimuli. As another result, perceptual learning was rather interval-specific, that is, training with 100-ms intervals did not substantially improve discrimination of 200-ms intervals. This finding is in line with the results of most previous results (Karmarkar and Buonomano 2003; Meegan et al. 2000; Nagarajan et al. 1998; Wright et al. 1997; but see Lapid et al. 2009).

In contrast to previous studies (Grondin and Ulrich 2011; Lapid et al. 2009), the present design not only allowed the assessment of transfer effects from audition to vision but also from vision to audition. As mentioned before, an asymmetric cross-modal transfer effect was revealed with this design, that is, perceptual learning transferred from audition to vision but not vice versa. This rather complex result pattern cannot be interpreted unequivocally in favour of one of the two opposing theoretical viewpoints, that is, dedicated and intrinsic models (Ivry and Schlerf 2008). Our major finding of a transfer of perceptual learning from the auditory to the visual modality is clearly inconsistent with the modality specificity of temporal processing implied in intrinsic models. The fact that perceptual learning did not or only very little transfer to intervals of other duration (see also Wright et al. 1997; Karmarkar and Buonomano 2003), however, argues against the existence of a centralized, interval-independent internal clock (e.g. the pacemaker-accumulator model). This interval specificity of perceptual learning is most consistent with multiple clock accounts (e.g. Buhusi and Meck 2009). Such models, however, cannot easily explain why perceptual learning transferred from the auditory to the visual modality but not vice versa.

This asymmetric transfer effect is consistent with the notion that the auditory modality has a dominant status in processing of temporal information. For example, it is a common finding that audition is more precise than vision in temporal perception (e.g. Penney et al. 2000; Wearden et al. 1998). Furthermore, audition dominates vision in many situations, for example, when visual and auditory temporal information conflict (Welch et al. 1986) or in sensorimotor coordination (Repp and Penel 2002). In a study of Guttman et al. (2005), incongruent auditory stimulation interfered with the ability to track a visual rhythm, but visual stimulation did not interfere with auditory tracking. Kanai et al. (2011) found that transcranial magnetic stimulation of the auditory cortex impaired time estimation of auditory and visual stimuli to the same

degree. Guttman et al. and Kanai et al. therefore suggested that time is primarily encoded in the auditory system and in temporal discrimination tasks, visual input is automatically transformed into an auditory representation.

Our finding of an asymmetric transfer from auditory training to the visual modality appears consistent with this cross-modal encoding hypothesis. If timing of visual stimuli, however, recruits the timing mechanism hosted by the auditory modality, then one would expect that even training within the visual modality improves the efficiency of this mechanism. Hence, perceptual learning with visual stimuli should also transfer to the auditory modality. Since this was not the case, one must assume that the efficiency of the timing mechanism can only be improved by training with auditory material. As a consequence, training effects with visual stimuli would mainly represent improvements of the cross-modal encoding mechanism.

Alternatively, the asymmetric transfer effect could also reflect memory-related processes. For example, within the attentional gate model by Zakay and Block (1997), one might assume that repeated presentations of a specific auditory interval form an especially stable amodal internal reference against each further interval in a given trial is compared. Accordingly, training with auditory stimuli improves not only auditory discrimination but also discrimination of stimuli within other modalities. In contrast, the amodal reference emerging from visual stimuli might—even after extensive training—be rather noisy. If this post-training reference is still noisier than the reference emerging from auditory stimuli, auditory discrimination does not benefit from visual training. An additional implication of this explanation is that cross-modal transfer should be interval-specific as is usually observed. Therefore, asymmetric cross-modal transfer effects might reflect differences in the formation of an internal standard (see Lapid et al. 2008) in the reference memory component of the attentional gate model.

In the present article, we investigated transfer effects between the auditory and the visual modality. As mentioned in the Introduction, previous research has shown such transfer effects from the somatosensory to the auditory modality (Nagarajan et al. 1998). One might consider this as evidence against the notion of auditory dominance in temporal processing. However, Nagarajan et al. did not examine potential transfer effects from the auditory to the somatosensory modality. It is conceivable that this transfer effect would be larger than the one from the somatosensory to the auditory modality observed by these authors. If so, such a finding would further strengthen the notion of auditory dominance. Future research is needed to address this issue.

In conclusion, the present study demonstrates cross-modal transfer of perceptual learning in temporal discrimination of sub-second intervals. This transfer was

asymmetric, that is, it occurred from vision to audition but not from audition to vision. This finding represents a novel illustration of auditory dominance in the processing of time and is consistent with the notion that time is primarily encoded in the auditory system.

Acknowledgments This study was supported by the Deutsche Forschungsgemeinschaft (UL 116/12-1). We thank Teresa Birngruber and Linda Idelberger for assistance in data acquisition.

References

- Banai K, Ortiz JA, Oppenheimer JD, Wright BA (2010) Learning two things at once: differential constraints on the acquisition and consolidation of perceptual learning. *Neuroscience* 165:436–444
- Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10:433–436
- Buhusi CV, Meck WH (2009) Relativity theory and time perception: single or multiple clocks? *PLoS One* 4:e6268
- Creelman CD (1962) Human discrimination of auditory duration. *J Acoust Soc Am* 34:582–593
- Fahle M, Edelman S (1993) Long-term learning in vernier acuity: effects of stimulus orientation, range and of feedback. *Vis Res* 33:397–412
- Gibbon J, Church RM, Meck WH (1984) Scalar timing in memory. *Ann NY Acad Sci* 423:52–77
- Grondin S (1993) Duration discrimination of empty and filled intervals marked by auditory and visual signals. *Percept Psychophys* 54:383–394
- Grondin S, Rammsayer T (2003) Variable foreperiods and temporal discrimination. *Q J Exp Psychol* 56A:731–765
- 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
- Guttman SE, Gilroy LA, Blake R (2005) Hearing what the eyes see: auditory encoding of visual temporal sequences. *Psych Sci* 16:228–235
- Ivry RB, Schlerf JE (2008) Dedicated and intrinsic models of time perception. *Trends Cogn Sci* 12:273–280
- 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
- Karmarkar UR, Buonomano DV (2003) Temporal specificity of perceptual learning in an auditory discrimination task. *Learn Mem* 10:141–147
- Karmarkar UR, Buonomano DV (2007) Timing in the absence of clocks: encoding time in neural network states. *Neuron* 53:427–438
- Lapid E, Rammsayer T, Ulrich R (2008) On estimating the difference limen in duration discrimination tasks: a comparison of the 2AFC and the reminder task. *Percept Psychophys* 70:291–305
- 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
- Meegan DV, Aslin RN, Jacobs RA (2000) Motor timing learned without motor training. *Nat Neurosci* 3:860–862
- Nagarajan SS, Blake DT, Wright BA, Byl N, Merzenich MM (1998) Practice-related improvements in somatosensory interval discrimination are temporally specific but generalize across skin location, hemisphere, and modality. *J Neurosci* 18:1559–1570
- Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* 10:437–442
- Penney TB, Gibbon J, Meck WH (2000) Differential effects of auditory and visual signals on clock speed and temporal memory. *J Exp Psychol Hum Percept Perform* 26:1770–1787
- Rammsayer TH (1994) Effects of practice and signal energy on duration discrimination of brief auditory intervals. *Percept Psychophys* 55:454–464
- Repp BH, Penel A (2002) Auditory dominance in temporal processing: new evidence from synchronization with simultaneous visual and auditory sequences. *J Exp Psychol Hum Percept Perform* 28:1085–1099
- Shiu L, Pashler H (1992) Improvement in line orientation discrimination is retinally local but dependent on cognitive set. *Percept Psychophys* 52:582–588
- Treisman M, Faulkner A, Naish PLN, Brogan D (1990) The internal clock: evidence for a temporal oscillator underlying time perception with some estimates of its characteristic frequency. *Percept* 19:705–743
- Ulrich R, Nitschke J, Rammsayer T (2006) Crossmodal temporal discrimination: assessing the predictions of a general pacemaker-counter model. *Percept Psychophys* 68:1140–1152
- Warm JS, Stutz RM, Vassolo PA (1975) Intermodal transfer in temporal discrimination. *Percept Psychophys* 18:281–286
- Wearden JH, Edwards H, Fakhri M, Percival A (1998) Why ‘‘sounds are judged longer than lights’’: application of a model of the internal clock in humans. *Q J Exp Psychol* 51B:97–120
- Welch RB, DuttonHurt LD, Warren DH (1986) Contributions of audition and vision to temporal rate perception. *Percept Psychophys* 39:294–300
- 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, Block RA (1997) Temporal cognition. *Curr Dir Psychol Sci* 6:12–16