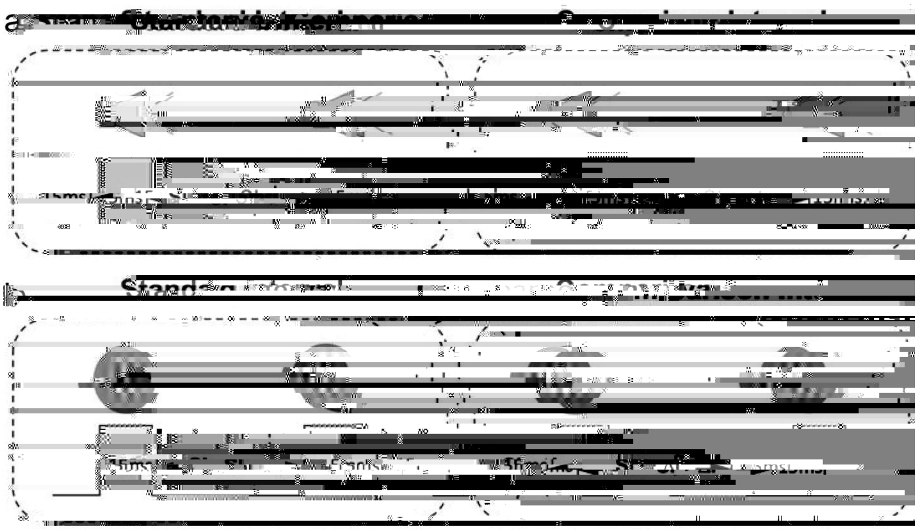
A supramodal and conceptual representation of subsecond time revealed with perceptual learning of temporal interval discrimination

Ying-Zi Xiong[1,2], Shu-Chen Guan[1,2] & Cong Yu[1]✉

Subsecond time perception has been frequently attributed to modality-specific timing mechanisms that would predict no cross-modal transfer of temporal perceptual learning. In fact, perceptual learning of temporal interval discrimination (TID) reportedly shows either no cross-modal transfer,



Illustrations of auditory and visual TID trials. ( ) An auditory TID trial. The standard stimuli were two 15-ms tone pips separated by a 100 ms interval, and the comparison stimuli were the same two tone pips separated by a 100 + Δt ms interval. In a given trial, the standard and comparison stimuli were presented in random order with a 900 ms time gap. ( ) A visual TID trial. The same as the auditory TID trial except that the tone-pips were replaced with Gabor patches.

The goal of this study is to demonstrate mutual and complete transfer of visual and auditory TID learning, so as to prove a supramodal subsecond time representation. Our previous perceptual learning studies have shown that various forms of specificities are not necessarily innate properties of perceptual learning as commonly believed, and can be eliminated with a double-training procedure[21–23]. In contrast to conventional training in which only the task of interest is practiced, double training consists of two training tasks. The primary training task in the current context would be TID in one sense (e.g., audition), and the secondary training task would be a functionally orthogonal one, such as contrast discrimination, in a new sense (e.g., vision). Here in the contrast discrimination task, the two Gabor gratings in a two-alternative forced-choice trial would mostly have near-threshold contrast differences and be presented at the same temporal interval as in the primary task, so that the participants would receive exposure to the temporal interval passively, but their attention is directed to near-threshold contrast discrimination to prevent potential temporal learning with the secondary task. The secondary task thus may activate sensory neurons representing the temporal interval in the new sense, so that the potential supramodal TID learning could functionally connect to temporal inputs from the new sense to improve TID performance. Double training has successfully enabled learning transfer of various visual discrimination tasks to untrained retinal location, orientation, motion direction, etc.[21,22,24–26]. It also succeeded in transferring auditory[27] and visuomotor learning[28,29].

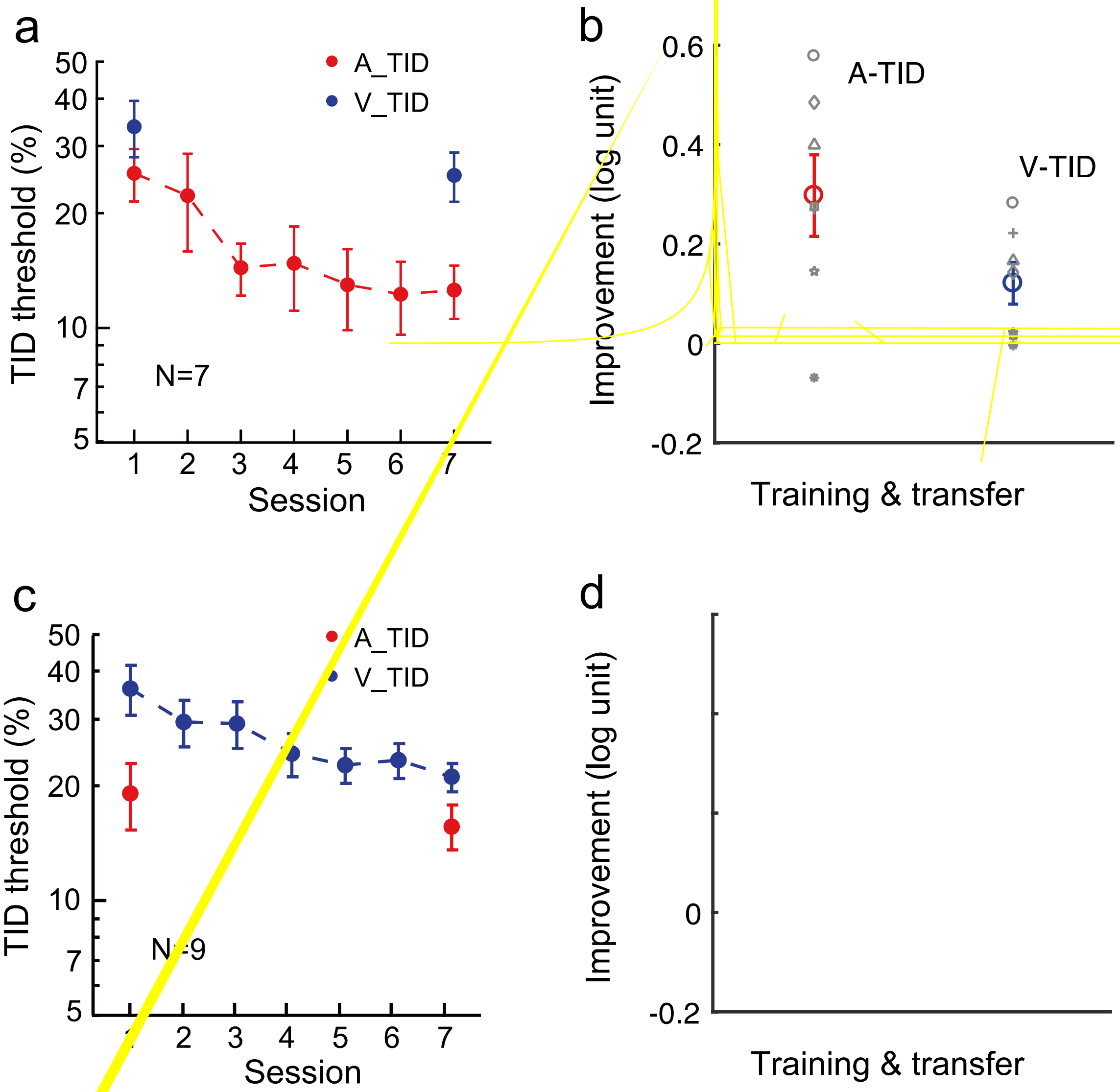
Most relevant to the current study is our recent report that perceptual learning of tactile orientation discrimination can transfer completely to visual orientation discrimination after double training, even if no transfer was evident with conventional single training[30]. These results are interpreted as evidence for a supramodal representation of stimulus orientation. Moreover, since the tactile orientation threshold is about three times as high as the visual orientation threshold, learning transfer is possible only if the supramodal representation is abstract and conceptual, independent of the original modality precision of sensory inputs[30,31]. Following the same reasoning, here we hypothesized that if perceptual learning of auditory and visual TID, which also differ in precision, could transfer mutually and completely with double training, we would also have evidence for a supramodal representation of subsecond time at a conceptual level.

Results

Baselines: asymmetric learning transfer between auditory and visual TID with conventional single training.

We first measured the cross-modal transfer of TID learning between audition and vision with conventional single training, which established baselines for later double training experiments. One group of participants ($N = 7$) practiced auditory TID (auditory single-training group), and a second group ($N = 9$) practiced visual TID (visual single-training group), both with the 100-ms standard interval.

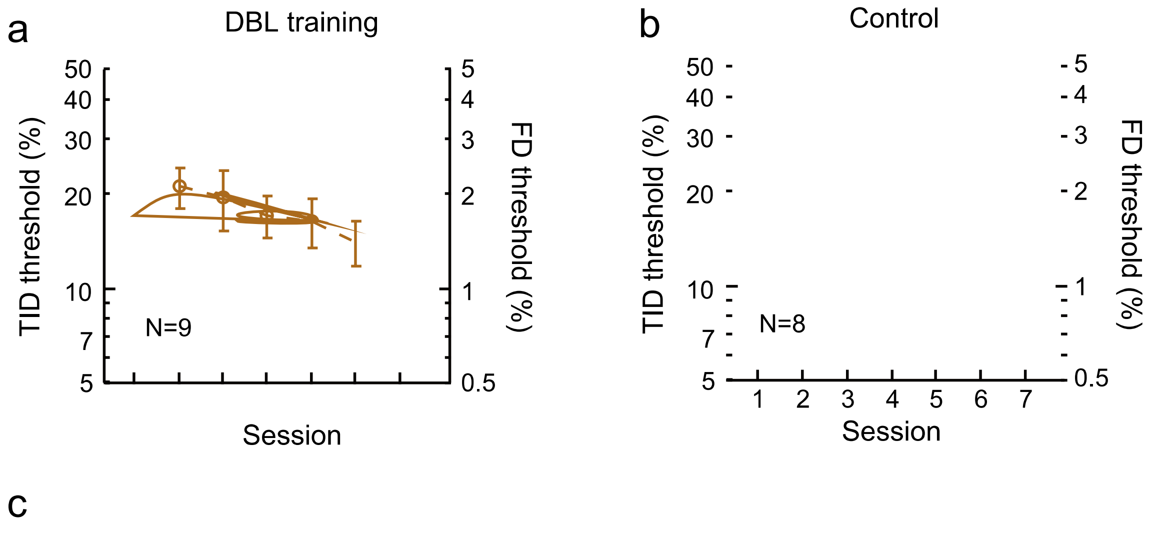
For the auditory single-training group, training reduced auditory TID threshold by 0.30 ± 0.08 log units ($t_6 = 3.63$, $p = 0.011$, Cohen's d = 1.37). The same training also improved visual TID at the same 100-ms interval, reducing visual TID threshold by 0.12 ± 0.04 log units ($t_6 = 3.87$, $p = 0.029$, Cohen's d = 1.08) (Fig. 2a, b). However, for the visual single-training group, although training improved visual TID by 0.20 ± 0.05 log units ($t_8 = 3.81$, $p = 0.005$, Cohen's d = 1.27), the learning did not transfer to auditory TID at the same interval (by 0.05 ± 0.05 log units; $t_8 = 1.04$, $p = 0.33$, Cohen's d = 0.35) (Fig. 2c, d). These results thus confirmed previous reports of asymmetric audition-to-vision transfer of TID learning with conventional single training[8,9]. Here the visual TID improvement through auditory TID training (V_TID in Fig. 2b) was about 60% of that through direct visual TID training



(V_TID in Fig. 2d), suggesting that auditory TID training might have not [illegible]aximized the visual TID performance in these observers. In other words, the audition-to-vision learnin[illegible]sfer was partial.

Double training: complete vision to auditio[illegible]er of TID learning. Next, we examined whether visual TID learning could transfer to a[illegible]D with double training. Nine participants practiced visual TID at a 100-ms interval. They also [illegible]posure to the auditory 100-ms interval by practicing an orthogonal tone frequency discrimina[illegible]he same interval. This double training improved visual TID by 0.21 ± 0.03 log units ($t_8 = 6.54$ [illegible]hen's d = 2.18) and tone frequency discrimination by 0.17 ± 0.05 log units ($t_8 = 3.44$, $p = 0.009$ [illegible].15) (Fig. 3a, c). Importantly, auditory TID at the same interval also showed an improveme[illegible] log units ($t_8 = 5.92$, $p < 0.001$, Cohen's d = 1.97) (Fig. 3c), which was not significantly di er[illegible].29 log-unit improvement with direct auditory TID training in the auditory single-trainin[illegible]) ($t_{14} = 0.63$, $p = 0.54$, Cohen's d = 0.31). Therefore, auditory TID appeared to have maximiz[illegible]D training and tone frequency discrimination training were coupled in double training [illegible]a ected by visual TID training alone (Fig. 2c, d).

[illegible]he possibility that the auditory TID improvement was simply a result of tone frequency dis[illegible] training, we had a control group ($N = 8$) only practice tone frequency discrimination at a 100-ms [illegible]al. The practice improved tone frequency discrimination by 0.17 ± 0.05 log units ($t_7 = 3.27$, $p = 0.014$, Cohen's [illegible] = 1.16), but it failed to improve auditory TID at the same interval (by − 0.03 ± 0.07 log units; $t_7 = -0.43$, p = 0.68, Figure 2



Cohen's d = −0.15, Fig. 3b, c). Taken together, the double training results and control data suggested that double training enabled full learning transfer from visual TID to auditory TID, in spite of the insignificant transfer in the single-training condition (Fig. 2c, d).

To reduce Type-I errors in our data analysis, a between-subject ANOVA compared auditory TID improvements among the three training conditions, i.e. single visual TID training, current double training, and tone frequency discrimination training. The ANOVA outputs suggested a significant main effect of training condition ($F_{2, 24} = 7.70$, $p = 0.003$, $\eta^2 = 0.39$). Further contrast analysis showed that the auditory TID improvement after double training was significantly higher than the improvement after single visual TID training ($t_{24} = 2.60$, $p = 0.016$) and the improvement after tone frequency discrimination training ($t_{26} = 2.69$, $p = 0.012$).

Double training: complete audition to vision transfer of TID learning. Earlier we suggested that visual TID improvement after auditory TID training was approximately 60% of that after direct visual TID training (Fig. 2b, d). Here we examined whether double training could lead to complete audition-to-vision TID learning transfer. Eight new participants practiced auditory TID and visual contrast discrimination, both at a 100-ms interval, in alternating blocks of trials in the same training sessions. Training improved auditory TID by

0.29 ± 0.04 log units ($t_7 = 6.55$, $p < 0.001$, Cohen's d = 2.32) and visual contrast discrimination by 0.37 ± 0.20 in d' ($t_7 = 2.06$, $p = 0.058$, Cohen's d = 0.73), as well as visual TID at the same interval by 0.26 ± 0.03 log units ($t_7 = 8.21$, $p < 0.001$, Cohen's d = 2.90) (Fig. 4a, c). The visual TID improvement did not differ significantly from the 0.20 log-unit improvement through direct visual TID training (Fig. 2c) ($t_{15} = 1.03$, $p = 0.31$, Cohen's d = 0.50), suggesting that the visual TID performance had maximized after double training.

Again, a control group of participants ($N = 8$) practiced visual contrast discrimination only, which improved contrast threshold by 0.84 ± 0.15 in d' ($t_7 = 4.74$, $p < 0.001$, Cohen's d = 1.68). But this practice had no significant impact on visual TID at the same 100-ms interval (by 0.04 ± 0.03 log units; $t_7 = 1.13$, $p = 0.30$, Cohen's d = 0.30, Fig. 4b, c), excluding the possibility that contrast discrimination training per se was responsible for above visual TID learning after double training. Here the visual pretraining threshold (V_TID) appeared to be lower than that with the double training group, which was mainly due to one participant who showed very low pre-training threshold at 11.8%. The pre-training V-TID thresholds were not significantly different from each other ($p = 0.146$, after single auditory TID training ($t_{20} = 2.74$, $p = 0.013$) and from the improvement after contrast discrimination training ($t_{20} = 3.23$, $p = 0.004$), confirming that double training induced more audition-to-vision TID learning transfer than auditory TID training alone, and that the double training effect could not be accounted for by visual contrast discrimination training.

Discussion

In this study we demonstrate mutual and complete cross-modal transfer of auditory and visual TID learning with double training, regardless of the difference in timing precisions (thresholds) between two senses, as well as the asymmetric audition-to-vision transfer of TID learning with conventional (single) training. These data thus provide direct support for a supramodal representation of subsecond time that can be improved through perceptual learning. Our results are consistent with previous reports which have also suggested supramodal subsecond time representation, on the basis of computer simulation[12], structure equation modeling of experimental data[14], and more direct crossmodal interference of duration judgments[13] and EEG data[11]. Evidence for a supramodal representation of subsecond time is in line with hypotheses of a dedicated central clock[1–3] that participates in subsecond time perception, although these hypotheses do not necessarily contradict the roles of distributed timing mechanisms[14].

The auditory and visual subsecond time information differs in not only modality origin, but also precision (the auditory TID threshold is approximately half the visual TID threshold, Figs. 2, 3, 4). Therefore, the double training results suggest complete cross-modal as well as cross-precision TID learning transfer. The cross-precision learning transfer would suggest that the time inputs from different modalities are represented equally at a supramodal level, which could be achieved through abstraction or standardization of the time inputs by their respective precisions (i.e., standard deviations). It is in this sense that we interpret the cross-modal TID learning transfer data as indications of not only supramodal, but also conceptual, representation of subsecond time. It is worth mentioning that the cross-modal TID learning transfer ma dal repudiae ptug39809(g)12.5(r(g)1.3(e in3d)-3 rate of 160 Hz. The luminance of the monitor was linearized by an 8-bit look-up table, with a mean luminance of 43.5 cd/m^2. A chin-and-head rest stabilized the head of the observer.

Stimuli and procedures. The auditory stimuli were two 15-ms tone pips separated by a 100 ms standard temporal interval (Fig. 1a). Each tone contained a 5-ms cosine ramp at each end, and was fixed at 1 kHz and 86 dB SPL. The visual stimuli were two 15-ms Gabor gratings, also separated by a 100 ms interval (Fig. 1b). Each Gabor had a fixed orientation (vertical), spatial frequency (1 cycle/deg), and contrast (100%). The length of the interval was the difference between the offset of the first stimulus and the onset of the second stimulus. We used 100 ms as the standard temporal interval because previous studies had shown clear evidence for significant TID learning and asymmetric audition-to-vision learning transfer at this interval[8].

The TID threshold was measured with a method of constant stimuli. In each forced-choice trial, a visual fixation was first centered on the computer screen for 300 ms, then two pairs of stimuli, one with a standard interval (100 ms) and the other with a comparison interval (100 ms + Δt), were subsequently presented in random order with a 900-ms time gap. The participants pressed the left or right arrow to indicate whether the first or the second pair of stimuli had a longer interval. A happy or sad cartoon face was shown on the screen after each response to indicate a correct or wrong response. A blank screen was presented before the next trial for a random duration (500-1000 ms). The Δt was set at 6 levels for each condition (auditory TID: ± 20.1, ± 13.4, ± 6.7 ms; visual TID: ± 33.5, ± 20.1, ± 6.7 ms), and the intervals between stimulus levels were increased if necessary to ensure a sufficient range of correct rates. Each level was repeated 10 times in a block of 60 trials, for a total of 5 blocks.

The psychometric function was fitted with $P = \frac{1}{1+e^{(-k)*(\Delta t-\Delta t_0)}}$, where P was the rate of reporting the comparison interval being longer at each Δt, k was the slope, and Δt_0 was the point of subjective equivalence. The TID threshold was equal to half the interquartile range of the function: $\text{Threshold} = (\Delta t_{.75} - \Delta t_{.25})/2$.

The stimuli for tone frequency discrimination were the same as those for auditory temporal interval discrimination, except that the frequencies of two pairs of pips were changed while the temporal intervals were fixed at 100 ms. Two pairs of tone pips, one pair at a standard frequency of 1 kHz and the other at a higher comparison frequency (1 kHz + Δf), were presented subsequently in a random order in each trial. The participants pressed the left or right arrow to indicate whether the first or second pair of tone pips had a higher frequency. A happy or sad cartoon face was provided as feedback.

The tone frequency discrimination threshold was measured with a temporal 2AFC staircase procedure. The starting frequency difference (Δf) between the standard and comparison stimuli was 50%, which decreased by a factor of 2 after every correct response until the first incorrect response. Then the Δf was varied by a factor of 1.414 following a 3-down-1-up staircase rule for a 79% correct rate. Each staircase ended after 60 trials. The threshold was calculated as the mean of the last 40 trials.

The stimuli used for visual contrast discrimination were the same as those for visual temporal interval discrimination, except that the Gabor contrast was varied while the interval was fixed (100 ms). Only one pair of Gabors was presented in each trial. In 80% of the trials, the two Gabors had identical contrast, which randomized from 0.15 to 1. In the remaining 20% trials, the contrasts of two Gabors differed by 50%. The participants judged whether two Gabors had identical contrast. A happy or sad cartoon face was provided as feedback. The d' value was calculated to measure the contrast discrimination performance.

Each experiment consisted of a pre-training session, five training sessions, and a post-training session on separate days. The experiment was completed within 7–13 days, with inter-session gaps of no more than 2 days. Each single-training session consisted of 16 blocks of trials and lasted for approximately 1.5 h. Each double-training session consisted of 10 blocks of trials for the primary task and 10 blocks of trials for the secondary task in an alternating order, and lasted for approximately 2 h.

Sample size. The sample size was decided on the basis of a previous TID learning study that used similar stimuli (100 ms–1 kHz condition in Fig. 4, ref.[18]). In our study, learning and transfer involved comparisons between pre- to post-training thresholds in all experiments. To achieve 80% power at $p = 0.05$, for a similar effect size of Cohen's d = 1.34 in ref.[18] when comparing pre- and post-training thresholds, a sample size of 7 would be required. We used a sample size of 9 for each experiment, with consideration of potential dropout of participants.

Data analysis. The TID thresholds were log-transformed to achieve normal distributions (Shapiro–Wilk test before log-transformation: $p < 0.001$ for auditory and visual TID thresholds; Shapiro–Wilk test after log-transformation: $p = 0.28$ and 0.60 for corresponding TID thresholds). The amount of TID learning or transfer was then measured by the difference of pre- and post-training thresholds in log unit. Data were analyzed with JASP 0.14.1. A two-tailed one-sampled t-test was performed to examine whether a learning or transfer effect was different from 0, and a between-subject ANOVA with Bonferroni's correction was performed for multiple comparisons.

Data availability

Data are available at https://github.com/visionplusplu/ModalityLearning.

Received: 16 January 2022; Accepted: 10 June 2022

References

1. Treisman, M. Temporal discrimination and the indifference interval: implications for a model of the "internal clock". *Psychol. Monogr.* , 1–31 (1963).

2. Creelman, C. D. Human discrimination of auditory duration. *J. Acoust. Soc. Am.*, 582–593 (1962).
3. Ivry, R. B. & Schlerf, J. E. Dedicated and intrinsic models of time perception. *Trends Cogn. Sci.*, 273–280 (2008).
4. Paton, J. J. & Buonomano, D. V. The neural basis of timing: Distributed mechanisms for diverse functions. *Neuron*, 687–705 (2018).
5. Grondin, S. Timing and time perception: a review of recent behavioral and neuroscience findings and theoretical directions. *Atten. Percept. Psychophys.*, 561–582 (2010).
6. Burr, D., Tozzi, A. & Morrone, M. C. Neural mechanisms for timing visual events are spatially selective in real-world coordinates. *Nat. Neurosci.*, 423–425 (2007).
7. Johnston, A., Arnold, D. H. & Nishida, S. Spatially localized distortions of event time. *Curr. Biol.*, 472–479 (2006).
8. Bratzke, D., Seifried, T. & Ulrich, R. Perceptual learning in temporal discrimination: asymmetric cross-modal transfer from audition to vision. *Exp. Brain Res.*, 205–210 (2012).
9. McGovern, D. P., Astle, A. T., Clavin, S. L. & Newell, F. N. Task-specific transfer of perceptual learning across sensory modalities. *Curr. Biol.*, R20-21 (2016).
10. Lapid, E., Ulrich, R. & Rammsayer, T. Perceptual learning in auditory temporal discrimination: no evidence for a cross-modal transfer to the visual modality. *Psychon. Bull. Rev.*, 382–389 (2009).
11. Barne, L. C. *et al.* A common representation of time across visual and auditory modalities. *Neuropsychologia*, 223–232 (2018).
12. Bratzke, D. & Ulrich, R. Temporal reproduction within and across senses: testing the supramodal property of the pacemaker-counter model. *J. Exp. Psychol. Hum.*, 1218–1235 (2019).
13. Filippopoulos, P. C., Hallworth, P., Lee, S. & Wearden, J. H. Interference between auditory and visual duration judgements suggests a common code for time. *Psychol. Res.*, 708–715 (2013).
14. Stauffer, C. C., Haldemann, J., Troche, S. J. & Rammsayer, T. H. Auditory and visual temporal sensitivity: evidence for a hierarchical structure of modality-specific and modality-independent levels of temporal information processing. *Psychol. Res.*, 20–31 (2012).
15. Merchant, H., Harrington, D. L. & Meck, W. H. Neural basis of the perception and estimation of time. *Annu. Rev. Neurosci.*, 313–336 (2013).
16. Wiener, M., Matell, M. S. & Coslett, H. B. Multiple mechanisms for temporal processing. *Front Integr. Neurosci.*, 31 (2011).
17. Bueti, B. & Buonomano, D. V. Temporal perceptual learning. *Timing Time Percept.*, 261–289 (2014).
18. Wright, B. A., Buonomano, D. V., Mahncke, H. W. & Merzenich, M. M. Learning and generalization of auditory temporal-interval discrimination in humans. *J. Neurosci.*, 3956–3963 (1997).
19. Guttman, S. E., Gilroy, L. A. & Blake, R. Hearing what the eyes see: auditory encoding of visual temporal sequences. *Psychol. Sci.*, 228–235 (2005).
20. Kanai, R., Lloyd, H., Bueti, D. & Walsh, V. Modality-independent role of the primary auditory cortex in time estimation. *Exp. Brain Res.*, 465–471 (2011).
21. Xiao, L. Q. *et al.* Complete transfer of perceptual learning across retinal locations enabled by double training. *Curr. Biol.*, 1922–1926 (2008).
22. Zhang, J. Y. *et al.* Rule-based learning explains visual perceptual learning and its specificity and transfer. *J. Neurosci. Off. J. Soc. Neurosci.*, 12323–12328 (2010).
23. Xiong, Y. Z., Zhang, J. Y. & Yu, C. Bottom-up and top-down influences at untrained conditions determine perceptual learning specificity and transfer. *Elife*, 14614 (2016).
24. Wang, R., Cong, L. J. & Yu, C. The classical TDT perceptual learning is mostly temporal learning. *J. Vis.* (5), 1–9 (2013).
25. Zhang, J. Y. & Yang, Y. X. Perceptual learning of motion direction discrimination transfers to an opposite direction with TPE training. *Vis. Res.*, 93–98 (2014).
26. Mastropasqua, T., Galliussi, J., Pascucci, D. & Turatto, M. Location transfer of perceptual learning: passive stimulation and double training. *Vis. Res.*, 93–102 (2015).
27. Xiong, Y. Z., Tang, D. L., Zhang, Y. X. & Yu, C. Complete cross-frequency transfer of tone frequency learning after double training. *J. Exp. Psychol. Gen.*, 94–103 (2020).
28. Yin, C., Bi, Y., Yu, C. & Wei, K. Eliminating direction specificity in visuomotor learning. *J. Neurosci.*, 3839–3847 (2016).
29. Grzeczkowski, L., Cretenoud, A. F., Mast, F. W. & Herzog, M. H. Motor response specificity in perceptual learning and its release by double training. *J. Vis.*, 4 (2019).
30. Hu, D. Z., Wen, K., Chen, L. H. & Yu, C. Perceptual learning evidence for supramodal representation of stimulus orientation at a conceptual level. *Vision. Res.*, 120–128 (2021).
31. Wang, R. *et al.* Perceptual learning at a conceptual level. *J. Neurosci.*, 2238–2246 (2016).
32. Cong, L. J., Wang, R. J., Yu, C. & Zhang, J. Y. Perceptual learning of basic visual features remains task specific with Training-Plus-Exposure (TPE) protocols. *J. Vis.* (3), 1–9 (2016).
33. Nagarajan, S. S., Blake, D. T., Wright, B. A., Byl, N. & Merzenich, M. M. Practice-related improvements in somatosensory interval discrimination are temporally specific but generalize across skin location, hemisphere, and modality. *J. Neurosci.*, 1559–1570 (1998).
34. Zakay, D. & Block, R. A. Temporal cognition. *Curr. Dir. Psychol. Sci.*, 12–16 (1997).
35. Zhang, G. L., Cong, L. J., Song, Y. & Yu, C. ERP P1–N1 changes associated with Vernier perceptual learning and its location specificity and transfer. *J. Vis.* (4), 1–13 (2013).
36. Pelli, D. G. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.*, 437–442 (1997).

Acknowledgements

This research was supported by a Ministry of Science and Technology, China grant 2022ZD0204601, a Natural Science Foundation of China Grant 31230030, and funds from Center for Life Sciences, Peking University.

Author contributions

Y.-Z.X. and C.Y. designed the experiments. Y.-Z.X. and S.-C.G. conducted the experiments. Y.-Z.X., S.-C.G. and C.Y. analyzed the data and wrote the paper. C.Y. supervised the project.

Competing interest

The authors declare no competing interests.

Additional information

and requests for materials should be addressed to C.Y.

is available at www.nature.com/reprints.

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.