




OPEN

## Effects of stimulus and task structure on temporal perceptual learning

Rannie Xu , Russell M. Church, Yuka Sasaki & Takeo Watanabe

Our ability to discriminate temporal intervals can be improved with practice. This learning is generally thought to reflect an enhancement in the representation of a trained interval, which leads to interval-specific improvements in temporal discrimination. In the present study, we asked whether temporal learning is further constrained by context-specific factors dictated through the trained stimulus and task structure. Two groups of participants were trained using a single-interval auditory discrimination task over 5 days. Training intervals were either one of eight predetermined values (FI group), or random from trial to trial (RI group). Before and after the training period, we measured discrimination performance using an untrained two-interval temporal comparison task. Our results revealed a selective improvement in the FI group, but not the RI group. However, this learning did not generalize between the trained and untrained tasks. These results highlight the sensitivity of TPL to stimulus and task structure, suggesting that mechanisms of temporal learning rely on processes beyond changes in interval representation.

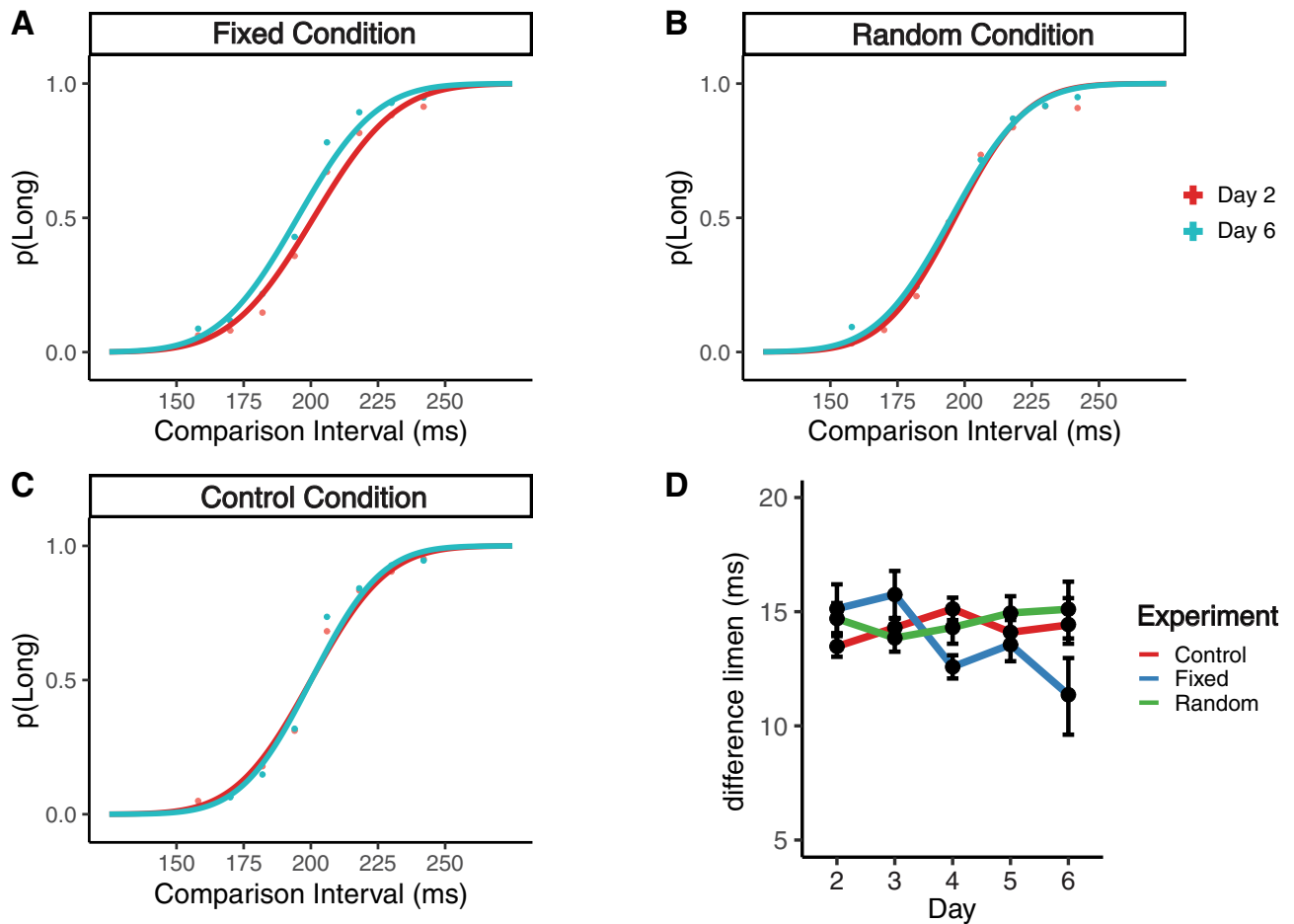
Temporal perceptual learning (TPL) refers to a training-related improvement in the detectability and discriminability of temporal intervals<sup>1</sup>. This type of learning can be long-lasting<sup>2</sup> and is generally thought to reflect a refinement in the representation of a learned interval<sup>3–6</sup>. Typically, TPL do not generalize to untrained conditions which differ from the trained temporal interval<sup>3–8</sup>. For example, Wright et al.<sup>9</sup> found that auditory discrimination training using a 100 ms tone selectively improved the discriminability of the trained interval, but this improvement did not generalize towards any neighboring conditions (e.g., 50 ms, 200 ms, 500 ms). Observation of temporally-specific learning is of key importance from a theoretical standpoint because it suggests that the mechanisms of TPL might involve duration-selective tuning mechanisms in sensory and/or association cortices of the brain<sup>10–12</sup>.

In addition to training-related refinements in temporal representation, TPL can similarly affect the way we respond to a temporal stimulus through non-temporal factors. For instance, optimizations in task-related processing (e.g., working memory capacity or decision bias) could change how sensory information is being interpreted in forming a behavioral choice; consequently, perceptual learning would rely on a strengthening in the relationship between an external stimulus and the optimal decision in a perceptual task<sup>13–15</sup>. In this view, TPL cannot exclusively reflect changes in the representation of a temporal interval, but also in the way we make decisions about that interval within a specific temporal task. Learning must therefore be *task-specific* and confined within the trained structure of information<sup>16</sup>—featural learning by itself is insufficient to account for the behavioral improvements observed with practice<sup>17–19</sup>.

Compared with perceptual learning in the visual modality (VPL), evidence of task-specificity in the temporal domain has been limited to a small number of studies exploring cross-modal effects of learning, and often producing mixed results<sup>4,20–23</sup>. Even within single studies, evidence of bi-directional transfer between the auditory and visual modalities have been inconclusive<sup>24,25</sup>. One interesting exception to this is a study by Meegan et al.<sup>20</sup> where training on a temporal discrimination task was found to improve motor production of the same interval. However, this type of task-independent transfer has not been documented within the same task type (e.g., strictly sensory or motor), and might be asymmetrical in nature. Therefore, it is unclear whether TPL is also accompanied by changes in task-specific processes.

The purpose of the present study is to abridge findings between VPL and TPL through a systematic investigation of task-specificity in temporal learning. If we assume that TPL shares a similar mechanism of learning to VPL in that performance improvement in perceptual tasks reflect changes to (1) a learned stimulus/feature and (2) the way we respond to the stimulus/feature within a task, we would expect TPL to be specific to the

Department of Cognitive, Linguistic and Psychological Sciences, Brown University, Providence 02912, USA. ✉email: ran\_xu@brown.edu



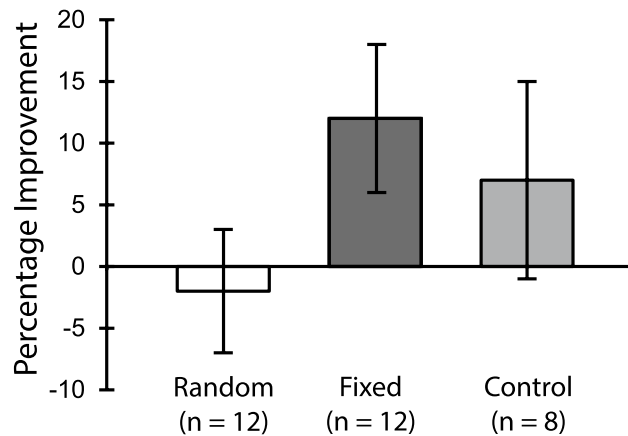
**Figure 1.** (a–c) Changes in the psychometric function in each group before and after training. (d) Changes in the difference limen (ms) across days. Error bars represent  $\pm$  SEM.

response strategies and demands within a learned task (i.e., no learning transfer to an unlearned task). To this end, we employed a similar pretest–training–posttest design as Meegan et al.<sup>20</sup>, where training involves extensive practice on a temporal discrimination task, and a pretest/posttest session which examines transfer of learning to an untrained task. Since we used the same temporal interval in both tasks, our design allows for an independent treatment of *interval-specific* versus *task-specific* learning. Further building off the results of Meegan et al.<sup>20</sup>, the use of two perceptual discrimination tasks in the present study is a more sensitive indicator of task-specificity in TPL, as it does not involve confounding factors such as motor efficiency or modality dominance in temporal processing<sup>24</sup>.

A second way in which we sought to investigate task-specificity in TPL concerns regularity in the stimulus distribution during training. One of the recent findings in VPL suggests that the degree of stimulus uncertainty or the ability to discern statistical regularities in the training stimulus might influence the type of decision strategy that is adopted by the viewer<sup>26–28</sup>. Consistent with this idea, we predict that TPL should be similarly sensitive to the stimulus structure during training since perceptual learning reflects a highly context-specific experience. As such, we also varied the type of stimulus input used during training. Half of the participants in our experiment received 8 fixed intervals (FI group) while the other half received random intervals (RI group) which varied from trial to trial with minimal certainty. We predict that the group receiving fixed input during training will show an improvement in discrimination performance, while the group trained using random intervals—as a result of lower statistical regularities in the training input—may not be able to adopt a response strategy appropriate to the task. In summary, we expect improvements in temporal discrimination to be strongly coupled to the trained structure of information, with no generalization beyond unlearned stimulus and task dimensions, akin to what is commonly demonstrated in VPL.

## Results

**Learning in the trained task.** To assess the amount of training-related change in performance between the FI and RI groups, we fit psychometric functions using the proportion of “long” responses at each comparison interval using the Quickpsy package<sup>29</sup> in R (Fig. 1). Based on the goodness-of-fit indicated by the coefficient of determination ( $r^2$ ), we excluded data from three outlying subjects.



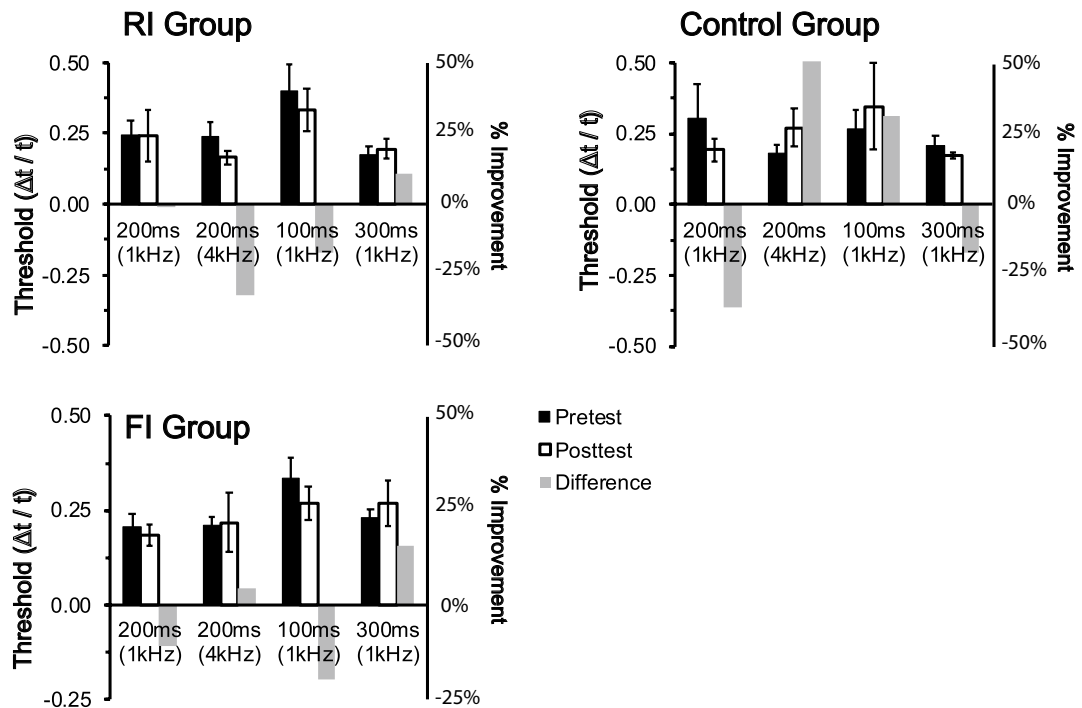
**Figure 2.** Percentage improvement on the trained temporal discrimination task. Percentage improvement is calculated from the change in difference limen between the first and last day of training (day 2–day 6), divided by an individual's performance on the first day of training (day 2): %Improvement = (day 2 – day 6)/day 2 × 100. This calculation measures the amount of improvement from baseline relative to an individual's initial performance on the task. Error bars indicate between-subjects ± SEM.

To investigate the effects of varying stimulus structure on TPL, we compared changes in timing ability of the RI and FI groups between the first and last day of training. We submitted the DL values to a 2 × 2 mixed ANOVA with group (FI/RI) as a between-subjects factor and session (day 2/day 6) as the within-subject factor. Overall, DL values decreased with training ( $F_{1,22} = 4.83$ ,  $p = 0.038$ ,  $\eta^2 = 0.09$ ; Day 2:  $M \pm SEM = 15.2 \pm 10.2$ ; Day 6:  $M \pm SEM = 12.9 \pm 6.6$ ), with no significant differences between RI and FI ( $F_{1,22} = 1.96$ ,  $p = 0.17$ ). However, we did observe a significant interaction effect between group and session ( $F_{1,22} = 4.41$ ,  $p = 0.047$ ,  $\eta^2 = 0.08$ ). A post-hoc pairwise t-test on the percentage change in DL between the FI group and the RI group revealed a significant selective improvement with training in the FI group ( $t_{11} = 2.17$ ,  $p = 0.041$ ), but not the RI group ( $t_{11} = 0.16$ ,  $p = 0.87$ ). To further qualify the degree of change in DLs within each group, we calculated a percent change in individual timing performance based on their initial performance on the first day of training (Fig. 2). For the FI group, which showed an average of 12% decrease in the DL across subjects, there was a significant learning effect as compared with the RI group with only an average of -2% change. These results are consistent with the idea that practice improves auditory bisection performance when a group receives a preset number of training intervals, but not when the test intervals belong to a random distribution. Our results suggest that TPL is similar to VPL in that learning is dependent on the input structure of information during training, and that the degree of improvement might be influenced by our ability to discern statistical regularities in the training stimulus<sup>30</sup>.

To further understand how statistical regularity contributed to the selective performance improvement in the FI group, we recruited nine additional participants for a control condition. In the RI group, every test interval was selected at random from a normal/gaussian distribution with a mean of 200 ms. This entails that a greater number of values are closer to 200 ms than any other value within that range. In contrast, the FI group always compared the reference against 1 of 8 fixed intervals with equal (uniform) probability. The use of a fixed number of values therefore ensures a greater degree of statistical regularity both through (1) the stimulus distribution, as well as (2) an overall smaller number of possible interval values—either/both of which might have contributed to the observed differences in learning between groups. If the latter is true, changing the stimulus distribution of the RI group from gaussian to uniform would have minimal effect on learning since the overall degree of certainty/variability of the training stimulus remains much lower than the FI group.

**Control group.** To address this possibility, we recruited a control group whereby all stimuli and parameters are identical to the RI condition, with the exception that test intervals are now drawn from a uniform—not gaussian—distribution. If TPL relies on regularities in the training stimuli, our manipulation would serve to equate the frequency of occurrence for all target intervals, leading to a similar behavioral improvement as observed in the FI group.

Once again, we fit individual psychometric functions for each session, and calculated a percentage change in DL. In the control group, we did not find a significant change in performance on the first and last day of training ( $t_7 = 1.22$ ,  $p = 0.26$ ), replicating results of the RI group. To see if there were any differences in learning between all three groups, we conducted a separate mixed model ANOVA with group (FI/RI/Control) and session (day 2/day 6) as factors. In this analysis, there was once again a significant, though reduced effect of training ( $F_{1,29} = 4.21$ ,  $p = 0.049$ ,  $\eta^2 = 0.06$ ), which likely reflects the lack of change in DL values in the RI and control groups. Also consistent with our prior results, we observed a significant group by session interaction ( $F_{2,29} = 3.57$ ,  $p = 0.041$ ,  $\eta^2 = 0.09$ ), and no differences between groups ( $F_{2,29} = 0.12$ ,  $p = 0.88$ ). These results again argue against the hypothesis that learning in the FI group is attributable to differences in the stimulus distribution (uniform vs. gaussian) during the training phase.



**Figure 3.** Discrimination thresholds (left) and percent improvement (right) for the untrained temporal comparison task. We did not observe a significant learning effect for the trained 200 ms interval between pretest and posttest sessions. Error bars represent  $\pm$  SEM.

**Generalization to the untrained task.** In addition to learning in the trained task, we also tested generalization of learning using an untrained discrimination task with the same reference interval. To compare discrimination performance across the 4 interval conditions (100 ms, 200 ms, 300 ms at 1 kHz and 200 ms at 4 kHz), Weber's Fractions (WF) were used (Fig. 3).

We then submitted the WFs to a  $3 \times 2 \times 4$  mixed model ANOVA with group as a between-subjects factor, and session and condition as within-subject factors. Our analysis revealed a significant main effect for condition ( $F_{3,87} = 4.01$ ,  $p = 0.038$ ,  $\eta^2 = 0.05$ ), with the 100 ms condition having significantly higher discrimination thresholds during both pretest ( $WF_{\text{pretest}} \pm \text{SEM} = 1.29 \pm 0.011$ ) and posttest ( $WF_{\text{posttest}} \pm \text{SEM} = 1.42 \pm 0.026$ ) sessions than all other interval conditions—a finding commonly reported in other TPL studies<sup>23,24,31</sup>. Importantly, we did not observe a significant effect of training ( $F_{6,87} = 0.88$ ,  $p = 0.45$ ), nor group by condition by session interaction ( $F_{6,87} = 1.39$ ,  $p = 0.22$ ). Together, these results suggest that given our current paradigm, training-related improvements on a temporal task does not generalize to an untrained task even if the interval used between the two tasks are identical. Of course, we must recognize that the two paradigms used in our study (i.e., temporal bisection and discrimination task) may not fully capture learning transfer across all types of temporal tasks, and that various task-related differences in structure, procedure, and design could potentially contribute to the null transfer observed in our experiment. However, our results provide an initial demonstration of task-specificity in TPL, possibly representing an additional constraint in the mechanisms of temporal learning.

## Discussion

Performance on a variety of perceptual and motor timing tasks can be improved with training<sup>3,9,20,32</sup>. This learning is largely interval-specific<sup>3–8</sup>, which motivates the view that TPL enhances the featural representation of a temporal interval. In contrast, relatively little is known about the degree to which non-temporal components of a task affect TPL. The present study explores the influence of stimulus- and task-specificity on temporal learning. In our experiment, we demonstrate that: (1) learning is sensitive to statistical regularities in the training stimulus—only the FI group, trained using a fixed number of interval values, showed a significant improvement in temporal discrimination; and (2) learning is specific to the response strategies of a trained task, with no generalization of learning to an untrained task.

We speculate that there are a few factors contributing to the selective improvement for the FI group, but not the RI group. Since the to-be-learned interval (i.e., 200 ms) was identical in both groups, this difference in learning can first and foremost be explained through differences in response strategy. For example, in lieu of comparing each test interval against the reference interval, participants in the FI group might have opted to learn a specific set of decision weights based on the optimal stimulus–response relationship. In other words, TPL reflects the learning of the most appropriate response given each of the 8 fixed values. With a finite number of possible values, the FI group can effectively adopt this strategy. By comparison, the RI (and control) group was not able to extract any meaningful relationships based on an infinite number of possible intervals and their

associated responses. This suggests that mechanisms of temporal learning might involve an optimization of connections between stable stimulus representations and relevant task decisions.

A related consequence of being exposed to greater variability in training input is the possibility of resulting differences in memory quality for the reference duration. Since the RI group was exposed to a greater number of test intervals during training, it is likely that their representation of the 200 ms reference is inherently noisier and more easily contaminated by additional input. As a result, degradations in memory may have contributed to the observed differences in learning across groups. This interpretation is again consistent with the stipulation that top-down processes are necessary in perceptual learning because it suggests that higher cognitive functions, such as attention and memory factors may influence perceptual processing at a very early stage of learning. With both considerations, our results suggest that TPL cannot rely *exclusively* on changes at the level of stimulus/interval representation, otherwise our manipulation of the comparison interval would not play a role in learning. Instead, we speculate that higher-level processes must also be involved in the learning process, and it is this set of task-specific strategies that also become improved with training.

More broadly, our finding that TPL is sensitive to regularities in the training stimulus is consistent with what is often reported as “roving effects” in VPL<sup>28</sup>. With roving, the use of interleaved visual stimulus or task structures can impair learning in a number of visual discrimination tasks<sup>26–28,33</sup>. This interference is suggested to result from activity in overlapping neural populations, and arising only when there is a high degree of similarity between a trained stimulus type and task<sup>34,35</sup>. In our study, the observation that learning was impaired for the RI group receiving interleaved training intervals therefore highlights a crucial similarity in the underlying learning mechanisms between temporal and visual perceptual learning. A recent finding by Parkosadze et al.<sup>36</sup> showed that perceptual training under roving conditions may be mitigated by a more extensive training regimen. Similarly, we might expect to see comparable learning outcomes for the RI group with a greater number of training trials.

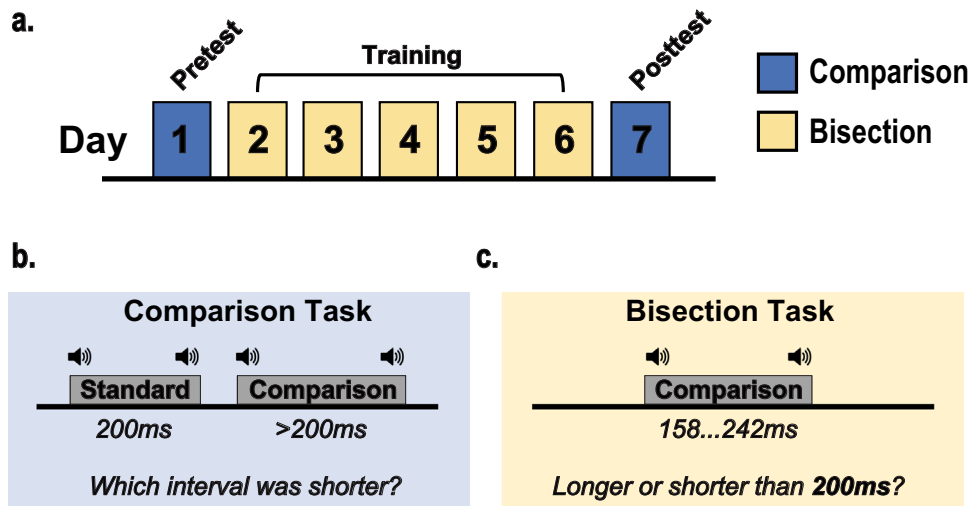
Another novel finding in our study is the specificity of learning within a trained temporal task. For the FI group, improvements in temporal discrimination in the practiced task was not transferred to the unlearned comparison task, even with the same reference interval. While our results differ from learning generalization reported in previous studies, one way to reconcile these discrepant findings is by considering the ease of generalization between trained and untrained conditions. Specifically, we speculate that learning generalization may be inversely proportional to task difficulty during training. In Buonomano et al.<sup>37</sup> for example, training using a short interstimulus interval (ISI) resulted in learning specificity for the trained condition, whereas the use of a longer ISI was able to benefit temporal discrimination in untrained conditions<sup>37</sup>. Similarly in Chen and Zhou<sup>38</sup>, transfer effects were robust for a visual Ternus display when trained using auditory/tactile intervals, but similar benefits were not observed in the visual modality alone. These results are consistent with the idea that learning generalization increases with training<sup>5</sup>, and that practice in a more difficult or cognitively demanding condition would impair transfer to untrained conditions. Conversely when training is conducted using easier (e.g., longer ISI) or modality-dominant (e.g., auditory) stimuli, learning transfer is typically observed. This suggests that differences in the relative ease in accessing unisensory temporal information can predict intermodal generalization patterns in untrained conditions.

Despite these considerations, we must acknowledge the possibility that the learning specificity reported in the present study could be attributable, at least in part, to differences in methodology used to estimate the discrimination threshold between the trained and untrained tasks. Since the testing phase involves an adaptive procedure (AP) rather than the method of constant stimuli (MCS), it is possible that our null transfer results reflect an insensitivity of one estimation method over the other, rather than an inability to generalize learning across tasks. For instance, since the AP involves the comparison of two interval durations, between which the difference is changed adaptively on every trial, this ever-changing difference could evoke the use of subtle irrelevant cues in interval discrimination<sup>39</sup> that is not available in MCS procedures. While both the AP and MCS are commonly used to estimate temporal discrimination performance in TPL research<sup>1</sup>, a direct comparison between the two methods have not been attempted, so we must recognize that inherent differences in threshold estimation may contribute to the learning specificity observed here.

In summary, our finding of task-based specificity is best understood as an additional constraint on the mechanisms of TPL. Based on several models in VPL, specificity in learning could reflect changes in a trained feature (i.e., reference interval) given a specific set of task- and context-specific demands<sup>40–43</sup>. Consistent with this hypothesis, our results extend our current understanding for the mechanisms of TPL by constraining interval-specific learning within a broader framework of task-related processes.

## Methods

**Participants.** Twenty-nine right-handed adults (16 females; mean age:  $23.8 \pm 3.9$  years) with normal or corrected-to-normal vision and hearing were recruited for participation in our experiment. We justified our sample size using the G\*Power software<sup>44</sup> with a medium effect size of 0.40 and a power estimate of 0.80. A total of 24 subjects were determined to be necessary for significance in a mixed-model ANOVA with this analysis (2 groups,  $F_{\text{critical}} = 1.05$ , non-centrality parameter = 184.32). Two participants were unable to complete the experiment due to scheduling conflicts, and 4 more outliers (3 from the FI/RI group, 1 from the control group) were excluded from analysis due to abnormally high thresholds during training ( $> 2.5SD$  from the group mean). For the control experiment, nine new participants were recruited from the same subject population (8 females; mean age:  $22.4 \pm 2.9$  years). Each session was held at the same time each day to avoid possible confounding effects of time of day on temporal processing<sup>45</sup>. Informed consent was obtained in writing from each subject prior to commencing the study and approved by the institutional review board (IRB) at Brown University.



**Figure 4.** (a) Experimental and task procedure. (b) Design for the untrained comparison task. Participants indicated which of two auditory tones was longer in duration (i.e., the comparison interval). Presentation order was randomized on every trial. (c) Design for the trained bisection task. Participants indicated whether the comparison interval was longer or shorter than 200 ms.

**Stimulus and apparatus.** Participants were seated in a sound-insulated room with dim lighting. All stimuli were generated and presented using MATLAB with Psychophysics Toolbox extensions, version 3.0.14<sup>46</sup> and presented on a ViewSonic—VA2226w monitor, measuring 20 × 14 in., with a refresh rate of 75 Hz and a viewing distance of approximately 38 cm. Auditory stimuli were presented at 86 dB SPL through noise-cancelling Sennheiser headphones and included a 5 ms on and off ramp. All responses were collected using a standard US keyboard.

**Procedure.** We used a standard pretest-training-posttest design over seven consecutive days (Fig. 4a). During the training phase (Days 2–6), each participant completed 2400 trials of a single-interval temporal discrimination task. We justified the number of practice trials under the guidance of previous studies which have demonstrated substantial learning effects using ~2500 trials<sup>6,20,24</sup>, and with as little as a single day of training with 900 trials<sup>32</sup>. At the beginning of each session, participants were presented with a 200 ms tone at 1 kHz and instructed to categorize subsequent intervals as being either “longer” or “shorter” than the reference (Fig. 4c). In the training phase, response feedback was provided immediately after every response by the participant. For the RI group, comparison intervals were drawn from a gaussian distribution with a mean of 200 ms and minimum and maximum bounds of 158 ms and 242 ms, respectively. In the FI group, 1 out of 8 predetermined interval values (158, 170, 182, 194, 206, 218, 230, 242 ms) were selected at random, with equal probability. Similar to the RI group, the mean and range of these intervals were 200 ms and 158–242 ms, respectively.

During the testing phase (Days 1 and 7), discrimination thresholds were measured using an adaptive staircase procedure<sup>47</sup> with four conditions: 100 ms, 200 ms, 300 ms at 1 kHz, and 200 ms at 4 kHz. The presentation order for each condition was pseudo-randomized according to the Latin Square design for each participant and kept constant between pretest and posttest sessions. The task involves a two-interval temporal comparison task (Fig. 4b) with a 3-down, 1-up adaptive staircase. Each training session consisted of 80 trials per test condition, and similar to estimation methods of previous studies<sup>9,48</sup>, we calculated discrimination thresholds based on the average of last 6 reversals within each block. On each trial, two auditory intervals were presented: a standard (t) and a comparison (t + Δt). Listeners must indicate which interval was the standard (i.e., shorter interval) by pressing either the “J” or “K” keys using their index and middle finger. The length of the standard and comparison was identical on the first trial of every block, forcing subjects to guess. The step size for the first five reversals is 5% of the standard and 1% thereafter. The inter-stimulus interval was jittered to minimize predictability between the offset of the first interval and the onset for the second interval. All methods and procedures were performed in accordance with the relevant guidelines and regulations set out by the IRB and human research protection program (HRPP) at Brown University.

**Data analysis.** For the training phase, the comparison intervals for the random condition were first divided into 8 equal bins such that the mean value for each bin corresponds to the comparison values in the fixed condition. It should be noted that the number of trials in each bin is different given the Gaussian distribution used in the RI condition. This discrepancy is addressed in our control group, which employs a uniform distribution similar to the fixed group (see “Results” section).

For each bin, we first calculated the proportion of “long” responses by dividing the number of times the participants responded “longer” to a given comparison value by the total number of responses within that bin (i.e., number of times the participant replied “short” plus number of times the participant replied “long”). Next,

we fit a psychometric function using the quickpsy package in R, allowing for lapses<sup>49</sup>. This function showed a monotonic increase with duration as participants generally responded “short” to the shortest duration and “long” to the longest durations. Based on these fitted functions, we then computed the just-noticeable-difference or the difference limen (DL) using the formula:

$$DL = \frac{x_{75\%} - x_{25\%}}{2}$$

where  $x$  represents the length of the comparison interval that yielded 25% and 75% “longer” responses, respectively. The DL therefore represents the steepness of the slope on the psychometric function and estimates the differential sensitivity of an observer during a timing task<sup>50</sup>. In this calculation, a smaller DL corresponds to a steeper psychometric function and a greater ability to discriminate between more similar intervals.

For the task used in the pretest/posttest sessions, we used a standard estimation technique for the discrimination threshold based on the average of the last 6 reversals in the staircase procedure. Since there were 4 different interval conditions, we then normalized each threshold measure by the reference duration. This is reported as the Weber’s Fraction (WF), with a smaller value corresponding to a higher degree of discriminability<sup>51</sup>:

$$WF = \frac{\Delta t}{\text{reference duration}}$$

where  $\Delta t$  represents the average difference between the comparison and reference intervals (in ms) on the last 6 reversals of the 1-up, 3-down staircase procedure. For instance, if the value of  $\Delta t$  is 40 ms for the 200 ms condition, it suggests that they are able to discriminate between 200 ms (reference) and 240 ms (comparison) with ~79% accuracy. We can then compute their WF by dividing the threshold by the interval condition (200 ms), resulting in a WF of 0.2. The use of the WF has the added advantage of comparing timing sensitivities across multiple conditions with different base reference values.

Received: 7 January 2020; Accepted: 17 December 2020

Published online: 12 January 2021

## References

- Buetti, D. & Buonomano, D. V. Temporal perceptual learning. *Timing Time Percept.* **2**, 261–289 (2014).
- Brown, D. J. & Proulx, M. J. Increased signal complexity improves the breadth of generalization in auditory perceptual learning. *Neural Plast.* <https://doi.org/10.1155/2013/879047> (2013).
- Karmarkar, U. R. & Buonomano, D. V. Temporal specificity of perceptual learning in an auditory discrimination task. *Learn. Mem.* **10**, 141–147 (2003).
- 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.* **18**, 1559–1570 (1998).
- Wright, B. A., Wilson, R. M. & Sabin, A. T. Generalization lags behind learning on an auditory perceptual task. *J. Neurosci.* **30**, 11635–11639 (2010).
- van Wassenhove, V. & Nagarajan, S. S. Auditory cortical plasticity in learning to discriminate modulation rate. *J. Neurosci.* **27**, 2663–2672 (2007).
- Buetti, D., Lasaponara, S., Cercignani, M. & Macaluso, E. Learning about time: plastic changes and interindividual brain differences. *Neuron* **75**, 725–737 (2012).
- Hayashi, M. J., van der Zwaag, W., Buetti, D. & Kanai, R. Representations of time in human frontoparietal cortex. *Commun. Biol.* **1**, 233 (2018).
- Wright, B. A., Buonomano, D. V., Mahncke, H. W. & Merzenich, M. M. Learning and generalization of auditory temporal-interval discrimination in humans. *J. Neurosci.* **17**, 3956–3963 (1997).
- Protopapa, F. *et al.* Chronotopic maps in human supplementary motor area. *PLoS Biol.* <https://doi.org/10.1371/journal.pbio.3000026> (2019).
- Hayashi, M. J. *et al.* Time adaptation shows duration selectivity in the human parietal cortex. *PLoS Biol.* **13**, e1002296 (2015).
- Harvey, B. M., Dumoulin, S. O., Fracasso, A. & Paul, J. M. A network of topographic maps in human association cortex hierarchically transforms visual timing-selective responses. *Curr. Biol.* <https://doi.org/10.1016/j.cub.2020.01.090> (2020).
- Dosher, B. A. & Lu, Z. L. Mechanisms of perceptual learning. in *Neurobiology of Attention* (2005). <https://doi.org/10.1016/B978-012375731-9/50082-3>
- Dosher, B. A. & Lu, Z.-L. Visual perceptual learning and models. *Annu. Rev. Vis. Sci.* <https://doi.org/10.1146/annurev-vision-102016-061249> (2017).
- Dosher, B. A. & Lu, Z.-L. Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proc. Natl. Acad. Sci.* <https://doi.org/10.1073/pnas.95.23.13988> (2002).
- Cohen, Y., Daikhin, L. & Ahissar, M. Perceptual learning is specific to the trained structure of information. *J. Cogn. Neurosci.* [https://doi.org/10.1162/jocn\\_a.00453](https://doi.org/10.1162/jocn_a.00453) (2013).
- Ahissar, M. & Hochstein, S. The reverse hierarchy theory of visual perceptual learning. *Trends Cogn. Sci.* **8**, 457–464 (2004).
- Ahissar, M., Nahum, M., Nelken, I. & Hochstein, S. Reverse hierarchies and sensory learning. *Philos. Trans. R. Soc. B Biol. Sci.* **364**, 285–299 (2009).
- Hochstein, S. & Ahissar, M. View from the top: hierarchies and reverse hierarchies. *Neuron* **36**, 791–804 (2002).
- Meegan, D. V., Aslin, R. N. & Jacobs, R. A. Motor timing learned without motor training. *Nat. Neurosci.* **3**, 860–862 (2000).
- Planetta, P. J. & Servos, P. Somatosensory temporal discrimination learning generalizes to motor interval production. *Brain Res.* <https://doi.org/10.1016/j.brainres.2008.07.081> (2008).
- Grondin, S. & Ulrich, R. Duration discrimination performance: No cross-modal transfer from audition to vision even after massive perceptual learning. in *Multidisciplinary Aspects of Time and Time Perception* 92–100 (2011). [https://doi.org/10.1007/978-3-642-21478-3\\_8](https://doi.org/10.1007/978-3-642-21478-3_8)
- 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.* **16**, 382–389 (2009).
- Bratzke, D., Seifried, T. & Ulrich, R. Perceptual learning in temporal discrimination: asymmetric cross-modal transfer from audition to vision. *Exp. Brain Res.* **221**, 205–210 (2012).

25. McGovern, D. P., Astle, A. T., Clavin, S. L. & Newell, F. N. Task-specific transfer of perceptual learning across sensory modalities. *Curr. Biol.* <https://doi.org/10.1016/j.cub.2015.11.048> (2016).
26. Adini, Y., Wilkonsky, A., Haspel, R., Tsodyks, M. & Sagi, D. Perceptual learning in contrast discrimination: the effect of contrast uncertainty. *J. Vis.* **4**, 2–2 (2004).
27. Kuai, S. G., Zhang, J. Y., Klein, S. A., Levi, D. M. & Yu, C. The essential role of stimulus temporal patterning in enabling perceptual learning. *Nat. Neurosci.* **8**, 1497 (2005).
28. Yu, C., Klein, S. A. & Levi, D. M. Perceptual learning in contrast discrimination and the (minimal) role of context. *J. Vis.* **4**, 169–182 (2004).
29. Linares, D. & López-Moliner, J. quickpsy: an R package to fit psychometric functions for multiple groups. *R J.* **8**, 122–131 (2019).
30. Banai, K. & Amitay, S. Stimulus uncertainty in auditory perceptual learning. *Vis. Res.* **4**, 2–2 (2012).
31. Warm, J. S., Stutz, R. M. & Vassolo, P. A. Intermodal transfer in temporal discrimination. *Percept. Psychophys.* <https://doi.org/10.3758/BF03199375> (1975).
32. Westheimer, G. Discrimination of short time intervals by the human observer. *Exp. Brain Res.* **129**, 121–126 (1999).
33. Otto, T. U., Herzog, M. H., Fahle, M. & Zhaoping, L. Perceptual learning with spatial uncertainties. *Vis. Res.* <https://doi.org/10.1016/j.visres.2006.03.021> (2006).
34. Zhang, J. Y. *et al.* Stimulus coding rules for perceptual learning. *PLoS Biol.* <https://doi.org/10.1371/journal.pbio.0060197> (2008).
35. Tartaglia, E. M., Aberg, K. C. & Herzog, M. H. Perceptual learning and roving: stimulus types and overlapping neural populations. *Vis. Res.* <https://doi.org/10.1016/j.visres.2009.02.013> (2009).
36. Parkosadze, K., Otto, T. U., Malania, M., Kezeli, A. & Herzog, M. H. Perceptual learning of bisection stimuli under roving: slow and largely specific. *J. Vis.* <https://doi.org/10.1167/8.1.5> (2008).
37. Buonomano, D. V., Bramen, J. & Khodadadifar, M. Influence of the interstimulus interval on temporal processing and learning: testing the state-dependent network model. *Philos. Trans. R. Soc. B Biol. Sci.* **364**, 1865–1873 (2009).
38. Chen, L. & Zhou, X. Fast transfer of crossmodal time interval training. *Exp. Brain Res.* <https://doi.org/10.1007/s00221-014-3877-1> (2014).
39. Xiong, Y. Z., Xie, X. Y. & Yu, C. Location and direction specificity in motion direction learning associated with a single-level method of constant stimuli. *Vis. Res.* <https://doi.org/10.1016/j.visres.2015.11.005> (2016).
40. Watanabe, T. & Sasaki, Y. Perceptual learning: toward a comprehensive theory. *Annu. Rev. Psychol.* **66**, 197–221 (2015).
41. Seitz, A. & Watanabe, T. A unified model for perceptual learning. *Trends Cogn. Sci.* **9**, 329–334 (2005).
42. Sasaki, Y., Nanez, J. E. & Watanabe, T. Advances in visual perceptual learning and plasticity. *Nat. Rev. Neurosci.* <https://doi.org/10.1038/nrn2737> (2010).
43. Sasaki, Y., Nánéz, J. E. & Watanabe, T. Recent progress in perceptual learning research. *Wiley Interdiscip. Rev. Cogn. Sci.* <https://doi.org/10.1002/wcs.1175> (2012).
44. Faul, F., Erdfelder, E., Lang, A. G. & Buchner, A. G\*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* **39**, 175–191 (2007).
45. Lustig, C. & Meck, W. H. Paying attention to time as one gets older. *Psychol. Sci.* **12**, 478–484 (2001).
46. Kleiner, M. *et al.* What's new in Psychtoolbox-3?. *Perception* **36**, S14 (2007).
47. Levitt, H. Transformed up-down methods in psychoacoustics. *J. Acoust. Soc. Am.* **49**, 467–477 (1971).
48. Wright, B. A., Sabin, A. T., Zhang, Y., Marrone, N. & Fitzgerald, M. B. Enhancing perceptual learning by combining practice with periods of additional sensory stimulation. *J. Neurosci.* <https://doi.org/10.1523/JNEUROSCI.0487-10.2010> (2010).
49. Wichmann, F. A. & Hill, N. J. The psychometric function: I. Fitting, sampling, and goodness of fit. *Percept. Psychophys.* <https://doi.org/10.3758/BF03194544> (2001).
50. Lapid, E., Ulrich, R. & Rammeyer, T. On estimating the difference limen in duration discrimination tasks: a comparison of the 2 AFC and the reminder task. *Percept. Psychophys.* **70**, 291–305 (2008).
51. Gibbon, J. Scalar expectancy theory and Weber's law in animal timing. *Psychol. Rev.* <https://doi.org/10.1037/0033-295X.84.3.279> (1977).

## Acknowledgements

This work was supported by the Edgar J. Marston Grant (EN464266) awarded to RMC, and the NIH R01EY027841 Grant for TW. We also thank Daniela Ravasio for assistance during data collection.

## Author contributions

R.X. and R.M.C. conceived of the research; R.X. completed the data collection and analyses under the guidance of Y.S. and T.W.; R.X. wrote the first draft of the manuscript; all authors participated in making revisions and approved the final version of the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to R.X.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021